

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 31, 2001, 08:52:06 ; Search time 23.87 Seconds
(without alignments)
37.238 Million cell updates/sec

Title: US-09-508-095-22

Perfect score: 60

Sequence: 1 ARRARYVWAAVG 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_1101.*

1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	60	100.0	12	20	AAW93886 Bifidobacterium bi
2	56	93.3	12	20	AAW93878 Bifidobacterium bi
3	56	93.3	13	20	AAW93880 Bifidobacterium bi
4	56	93.3	687	19	AAW71183 Mutant human lacto
5	56	93.3	688	19	AAW71182 Mutant human lacto
6	56	93.3	689	19	AAW71181 Mutant human lacto
7	56	93.3	690	19	AAW71180 Mutant human lacto
8	56	93.3	692	21	AAV58733 Human lactoferrin
9	56	93.3	692	22	AAW97382 Human lactoferrin
10	56	93.3	693	16	AAW85146 Lactoferrin protel
11	56	93.3	705	13	AAW22423 Human lactoferrin

12	56	93.3	705	20	AAV31152 Human lactoferrin
13	56	93.3	708	13	AAW22424 Human lactoferrin
14	56	93.3	708	20	AAV31153 Human lactoferrin
15	56	93.3	709	12	AAW12583 Lactoferrin. Homo
16	56	93.3	709	15	AAW45198 Human lactoferrin
17	56	93.3	709	15	AAW45199 Human lactoferrin
18	56	93.3	709	18	AAW21695 Human lactoferrin
19	56	93.3	709	19	AAW53879 Human lactoferrin
20	56	93.3	709	21	AAW03830 Human lactoferrin
21	56	93.3	709	21	AAW77577 Human lactoferrin
22	56	93.3	709	22	AAW36842 Protein encoded by
23	56	93.3	711	11	AAW08033 Human lactoferrin
24	56	93.3	711	14	AAW43653 Lactoferrin. Homo
25	56	93.3	711	17	AAW09342 Human lactoferrin
26	56	93.3	711	19	AAW57317 Human lactoferrin
27	56	93.3	711	19	AAW53880 Bovine alphaSI-cas
28	56	93.3	711	20	AAW86021 Human lactoferrin
29	56	93.3	711	21	AAW08182 Amino acid sequenc
30	56	93.3	711	21	AAW03831 Human lactoferrin
31	56	93.3	711	21	AAW77578 Human lactoferrin
32	56	93.3	711	22	AAW64828 Chronic hepatitis
33	56	93.3	711	22	AAW02341 Human lactoferrin
34	56	93.3	711	22	AAW36843 Human lactoferrin
35	49	81.7	685	12	AAW11663 Partial porcine la
36	49	81.7	685	12	AAW11664 Partial porcine hu
37	49	81.7	703	17	AAW09344 Porcine lactoferr
38	49	81.7	703	19	AAW57319 Porcine lactoferr
39	49	81.7	703	20	AAW86023 Porcine lactoferr
40	49	81.7	703	21	AAW08184 Porcine lactoferr
41	49	81.7	703	22	AAW02343 Amino acid sequenc
42	49	81.7	708	16	AAW71037 Draculin. Desmodu
43	45	75.0	689	21	AAW49270 Lactoferrin sequen
44	45	75.0	708	17	AAW09343 Bovine lactoferrin
45	45	75.0	708	19	AAW57318 Bovine lactoferrin

ALIGNMENTS

RESULT 1

AAW93886

ID AAW93886 standard; peptide; 12 AA.

XX AC AAW93886;

XX DT 25-JUN-1999 (first entry)

XX DE Bifidobacterium bifidus stimulating peptide 22.

XX DE Bifidogenic peptide; protease; treatment; microbe-related disease;

XX KW bacteria; fungi; yeast; protozoa; virus; mycoplasma; filaria; plasmodia;

XX KW infection; inflammation; microbial induced tumour; degenerative disorder;

XX KW diarrhoea; colic; oral microflora; intestinal microflora; caries;

XX KW vaginal microflora.

XX OS Bifidobacterium bifidus.

XX PN WO9914231-A2.

XX PD 25-MAR-1999.

XX PF 16-SEP-1998; 98WO-EP05899.

XX PR 11-FEB-1998; 98DE-1005385.

XX PR 16-SEP-1997; 97DE-1040604.

XX PA (FORS/) FORSMANN W.

XX PI Forssmann W, Liepke C, Zucht H;

XX PI WPI; 1999-244022/20.

XX PT Milk-derived peptides that stimulate Bifidobacterium bifidus

XX PS Claim 2; Page 3; 25pp; German.

XX CC This invention describes milk-derived bifidogenic peptides and their

XX CC active derivatives or fragments, and combinations of them produced by

XX CC chemical coupling. Such are produced from bovine or human milk by

XX CC treatment for 2 hr with proteases, then centrifuging to remove fat and

XX CC acidifying to pH 2 to precipitate proteins. The solution phase is then

XX CC subjected to reverse-phase HPLC, the fractions adjusted to salt content below

XX CC 25 mM (by dialysis or reverse-phase HPLC) and tested for activity by

XX CC growing Bifidobacterium bifidus and Escherichia coli in presence of the

XX CC fractions. Those fractions for which (Bw-B0)-(Ew-E0) is at least 0.15

XX CC 50% Elliker broth containing peptide at 0.2 mg/ml, B0 = germ count under

XX CC similar conditions in a peptide-free control, Ew = germ count after 16

XX CC hr culture of E. coli in 3 g/l tryptic broth containing peptide at

XX CC 0.2 mg/ml, E0 = germ count under similar conditions in a peptide-free

XX CC control. The peptides AAW93878-W93888 are used to treat microbe-related

XX CC diseases caused by bacteria, fungi, yeast, protozoa, viruses,

XX CC mycoplasma, filaria and plasmodia, e.g. infections, inflammation,

XX CC microbially induced tumours or degenerative disorders, diarrhoea, colic,

XX CC abnormalities in oral, intestinal or vaginal microflora, or caries.

XX SQ Sequence 12 AA;

Query Match 100.0%; Score 60; DB 20; Length 12;

Best Local Similarity 100.0%; Pred. No. 0.00031;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARRVRVVAAG 12

DB 1 arrrvvvaag 12

RESULT 2

AAW93878

ID AAW93878 standard; peptide; 12 AA.

XX AC AAW93878;

XX DT 25-JUN-1999 (first entry)

XX DE Bifidobacterium bifidus stimulating peptide 14.

XX KW Bifidogenic peptide; protease; treatment; microbe-related disease;

XX KW bacteria; fungi; yeast; protozoa; virus; mycoplasma; filaria; plasmodia;

XX KW infection; inflammation; microbial induced tumour; degenerative disorder;

XX KW diarrhoea; colic; oral microflora; intestinal microflora; caries;

XX KW vaginal microflora.

XX OS Bifidobacterium bifidus.

XX PN WO9914231-A2.

XX PD 25-MAR-1999.

XX PF 16-SEP-1998; 98WO-EP05899.

XX PR 11-FEB-1998; 98DE-1005385.

XX PR 16-SEP-1997; 97DE-1040604.

XX PA (FORS/) FORSMANN W.

XX PI Forssmann W, Liepke C, Zucht H;

XX DR WPI; 1999-244022/20.

XX PT Milk-derived peptides that stimulate Bifidobacterium bifidus

XX PS Claim 2; Page 3; 25pp; German.

CC This invention describes milk-derived bifidogenic peptides and their

CC active derivatives or fragments, and combinations of them produced by

CC chemical coupling. Such are produced from bovine or human milk by

CC treatment for 2 hr with proteases, then centrifuging to remove fat and

CC acidifying to pH 2 to precipitate proteins. The solution phase is then

CC subjected to reverse-phase HPLC, the fractions adjusted to salt content below

CC 25 mM (by dialysis or reverse-phase HPLC) and tested for activity by

CC growing Bifidobacterium bifidus and Escherichia coli in presence of the

CC fractions. Those fractions for which (Bw-B0)-(Ew-E0) is at least 0.15

CC 50% Elliker broth containing peptide at 0.2 mg/ml, B0 = germ count under

CC similar conditions in a peptide-free control, Ew = germ count after 16

CC hr culture of E. coli in 3 g/l tryptic broth containing peptide at

CC 0.2 mg/ml, E0 = germ count under similar conditions in a peptide-free

CC control. The peptides AAW93865-W93888 are used to treat microbe-related

CC diseases caused by bacteria, fungi, yeast, protozoa, viruses,

CC mycoplasma, filaria and plasmodia, e.g. infections, inflammation,

CC microbially induced tumours or degenerative disorders, diarrhoea, colic,

CC abnormalities in oral, intestinal or vaginal microflora, or caries.

XX SQ Sequence 12 AA;

Query Match 93.3%; Score 56; DB 20; Length 12;

Best Local Similarity 91.7%; Pred. No. 0.0015;

Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ARRVRVVAAG 12

DB 1 arrrvvvaag 12

RESULT 3

AAW93880

ID AAW93880 standard; peptide; 13 AA.

XX AC AAW93880;

XX DT 25-JUN-1999 (first entry)

XX DE Bifidobacterium bifidus stimulating peptide 16.

XX KW Bifidogenic peptide; protease; treatment; microbe-related disease;

XX KW bacteria; fungi; yeast; protozoa; virus; mycoplasma; filaria; plasmodia;

XX KW infection; inflammation; microbial induced tumour; degenerative disorder;

XX KW diarrhoea; colic; oral microflora; intestinal microflora; caries;

XX KW vaginal microflora.

XX OS Bifidobacterium bifidus.

XX PN WO9914231-A2.

XX PD 25-MAR-1999.

XX PF 16-SEP-1998; 98WO-EP05899.

XX PR 11-FEB-1998; 98DE-1005385.

XX PR 16-SEP-1997; 97DE-1040604.

XX PA (FORS/) FORSMANN W.

XX PI Forssmann W, Liepke C, Zucht H;

XX DR WPI; 1999-244022/20.

XX PT Milk-derived peptides that stimulate Bifidobacterium bifidus

XX PS Claim 2; Page 3; 25pp; German.

XX CC This invention describes milk-derived bifidogenic peptides and their

XX CC active derivatives or fragments, and combinations of them produced by

XX CC chemical coupling. Such are produced from bovine or human milk by

CC treatment for 2 hr with proteases, then centrifuging to remove fat and
 CC acidifying to pH 2 to precipitate proteins. The solution phase is then
 CC subjected to reverse-phase high-performance liquid chromatography (HPLC)
 CC and cation-exchange HPLC, the fractions adjusted to salt content below
 CC 25 mM (by dialysis or reverse-phase HPLC) and tested for activity by
 CC growing *Bifidobacterium bifidus* and *Escherichia coli* in presence of the
 CC fractions. Those fractions for which (Bw-B0)-(Ew-E0) is at least 0.15
 CC are selected where Bw = germ count after 16 hr culture of *B. bifidus* in
 CC 50% Elliker broth containing peptide at 0.2 mg/ml, B0 = germ count under
 CC similar conditions in a peptide-free control, Ew = germ count after 16
 CC hr culture of *E. coli* in 3 g/l tryptic broth containing peptide at
 CC 0.2 mg/ml, E0 = germ count under similar conditions in a peptide-free
 CC control. The peptides AAW93865-W93888 are used to treat microbe-related
 CC diseases caused by bacteria, fungi, yeast, protozoa, viruses,
 CC mycoplasma, filaria and plasmodia, e.g. infections, inflammation,
 CC microbially induced tumours or degenerative disorders, diarrhoea, colic,
 CC abnormalities in oral, intestinal or vaginal microflora, or caries.
 XX
 SQ Sequence 13 AA;

Query Match 93.3%; Score 56; DB 20; Length 13;
 Best Local Similarity 91.7%; Pred. No. 0.0016;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ARARVVVAAGV 12
 Db 1 arrarvvvcavg 12
 ||||| |||

RESULT 4
 AAW71183
 ID AAW71183 standard; protein; 687 AA.
 XX
 AC AAW71183;
 XX
 DT 27-OCT-1998 (first entry)
 XX
 DE Mutant human lactoferrin protein designated hLF-5N.
 XX
 KW Human; lactoferrin protein; variant; anti-inflammatory; anti-viral;
 KW anti-infective; coagulant; complement activation; inhibition;
 KW LPS mediated activation; myelopoiesis; growth promotion;
 KW intestinal epithelial cell; hydroxyl-radical formation;
 KW intestinal iron uptake; excretion.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN W09833509-A2.
 XX
 PD 06-AUG-1998.
 XX
 PF 02-FEB-1998; 98WO-IB00441.
 XX
 PR 02-FEB-1998; 98US-0017043.
 PR 03-FEB-1997; 97US-0036859.
 XX
 PA (PHAR-) PHARMING BV.
 XX
 PI Nuijens J, Van Berkel PHC;
 XX
 DR WPI; 1998-437164/37.
 XX
 PT Compositions containing human lactoferrin and variants - are used
 PT for treating e.g. anaemia, iron-storage disease, inflammation,
 PT tumours, rheumatoid arthritis, ulcerative colitis or infections
 XX
 PS Claim 2; Page -; 70pp; English.
 XX
 CC AAW71180-83 represent N-terminally truncated human lactoferrin
 CC proteins. The lactoferrin variants bind heparin with lower affinity than
 CC natural lactoferrin. The lactoferrin variants exhibit biological

CC activities such as anti-inflammatory, anti-viral and anti-infective
 CC activities as well as a pro- and anti-coagulant effects, modulation of
 CC complement activation, inhibition of LPS mediated activation of
 CC neutrophils, inhibition of myelopoiesis, regulation of transcription,
 CC growth promotion of intestinal epithelial cells, inhibition of
 CC hydroxyl-radical formation, and a role in intestinal iron uptake and
 CC excretion.
 CC note: this sequence does not appear in the specification; it was created
 CC using information provided.
 XX
 SQ Sequence 687 AA;

Query Match 93.3%; Score 56; DB 19; Length 687;
 Best Local Similarity 91.7%; Pred. No. 0.08;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ARARVVVAAGV 12
 Db 336 arrarvvvcavg 347
 ||||| |||

RESULT 5
 AAW71182
 ID AAW71182 standard; protein; 688 AA.
 XX
 AC AAW71182;
 XX
 DT 27-OCT-1998 (first entry)
 XX
 DE Mutant human lactoferrin protein designated hLF-4N.
 XX
 KW Human; lactoferrin protein; variant; anti-inflammatory; anti-viral;
 KW anti-infective; coagulant; complement activation; inhibition;
 KW LPS mediated activation; myelopoiesis; growth promotion;
 KW intestinal epithelial cell; hydroxyl-radical formation;
 KW intestinal iron uptake; excretion.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN W09833509-A2.
 XX
 PD 06-AUG-1998.
 XX
 PF 02-FEB-1998; 98WO-IB00441.
 XX
 PR 02-FEB-1998; 98US-0017043.
 PR 03-FEB-1997; 97US-0036859.
 XX
 PA (PHAR-) PHARMING BV.
 XX
 PI Nuijens J, Van Berkel PHC;
 XX
 DR WPI; 1998-437164/37.
 XX
 PT Compositions containing human lactoferrin and variants - are used
 PT for treating e.g. anaemia, iron-storage disease, inflammation,
 PT tumours, rheumatoid arthritis, ulcerative colitis or infections
 XX
 PS Claim 2; Page -; 70pp; English.
 XX
 CC AAW71180-83 represent N-terminally truncated human lactoferrin
 CC proteins. The lactoferrin variants bind heparin with lower affinity than
 CC natural lactoferrin. The lactoferrin variants exhibit biological
 CC activities such as anti-inflammatory, anti-viral and anti-infective
 CC activities as well as a pro- and anti-coagulant effects, modulation of
 CC complement activation, inhibition of LPS mediated activation of
 CC neutrophils, inhibition of myelopoiesis, regulation of transcription,
 CC growth promotion of intestinal epithelial cells, inhibition of
 CC hydroxyl-radical formation, and a role in intestinal iron uptake and
 CC excretion.
 CC note: this sequence does not appear in the specification; it was created

CC using information provided.

SQ Sequence 688 AA;

Query Match 93.3%; Score 56; DB 19; Length 688;

Best Local Similarity 91.7%; Pred. No. 0.08; Mismatches 1; Indels 0; Gaps 0;

QY 1 ARRVRVWAAVG 12

|||||||

Db 337 arrarvvwcavg 348

RESULT 6

AAW71181.
ID AAW71181 standard; protein; 689 AA.

AC AAW71181;

XX 27-OCT-1998 (first entry)

DE Mutant human lactoferrin protein designated hLF-3N.

XX Human; lactoferrin protein; variant; anti-inflammatory; anti-viral;
KW anti-infective; coagulant; complement activation; inhibition;
KW LPS mediated activation; myelopoiesis; growth promotion;
KW intestinal epithelial cell; hydroxyl-radical formation;
KW intestinal iron uptake; excretion.

XX Synthetic.

OS Homo sapiens.

XX WO9833509-A2.

XX 06-AUG-1998.

XX 02-FEB-1998; 98WO-IB00441.

XX 02-FEB-1998; 98US-0017043.

XX 03-FEB-1997; 97US-0036859.

XX (PHAR-) PHARMING BV.

XX Nuijens J, Van Berkel PHC;

XX WPI; 1998-437164/37.

XX Compositions containing human lactoferrin and variants - are used
PT for treating e.g. anaemia, iron-storage disease, inflammation,
PT tumours, rheumatoid arthritis, ulcerative colitis or infections

XX Claim 2; Page -; 70pp; English.

XX AAW71180-83 represent N-terminally truncated human lactoferrin
CC proteins. The lactoferrin variants bind heparin with lower affinity than
CC natural lactoferrin. The lactoferrin variants exhibit biological
CC activities such as anti-inflammatory, anti-viral and anti-infective
CC activities as well as a pro- and anti-coagulant effects, modulation of
CC complement activation, inhibition of LPS mediated activation of
CC neutrophils, inhibition of myelopoiesis, regulation of transcription,
CC growth promotion of intestinal epithelial cells, inhibition of
CC hydroxyl-radical formation, and a role in intestinal iron uptake and
CC excretion.

CC note: this sequence does not appear in the specification; it was created
CC using information provided.

XX Sequence 689 AA;

Query Match 93.3%; Score 56; DB 19; Length 689;

Best Local Similarity 91.7%; Pred. No. 0.08; Mismatches 1; Indels 0; Gaps 0;

Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ARRVRVWAAVG 12

|||||||

Db 338 arrarvvwcavg 349

RESULT 7

AAW71180

ID AAW71180 standard; protein; 690 AA.

XX AAW71180;

XX 27-OCT-1998 (first entry)

DE Mutant human lactoferrin protein designated hLF-2N.

XX Human; lactoferrin protein; variant; anti-inflammatory; anti-viral;
KW anti-infective; coagulant; complement activation; inhibition;
KW LPS mediated activation; myelopoiesis; growth promotion;
KW intestinal epithelial cell; hydroxyl-radical formation;
KW intestinal iron uptake; excretion.

XX Synthetic.

OS Homo sapiens.

XX WO9833509-A2.

XX 06-AUG-1998.

XX 02-FEB-1998; 98WO-IB00441.

XX 02-FEB-1998; 98US-0017043.

XX 03-FEB-1997; 97US-0036859.

XX (PHAR-) PHARMING BV.

XX Nuijens J, Van Berkel PHC;

XX WPI; 1998-437164/37.

XX Compositions containing human lactoferrin and variants - are used
PT for treating e.g. anaemia, iron-storage disease, inflammation,
PT tumours, rheumatoid arthritis, ulcerative colitis or infections

XX Claim 2; Page -; 70pp; English.

XX AAW71180-83 represent N-terminally truncated human lactoferrin
CC proteins. The lactoferrin variants bind heparin with lower affinity than
CC natural lactoferrin. The lactoferrin variants exhibit biological
CC activities such as anti-inflammatory, anti-viral and anti-infective
CC activities as well as a pro- and anti-coagulant effects, modulation of
CC complement activation, inhibition of LPS mediated activation of
CC neutrophils, inhibition of myelopoiesis, regulation of transcription,
CC growth promotion of intestinal epithelial cells, inhibition of
CC hydroxyl-radical formation, and a role in intestinal iron uptake and
CC excretion.

CC note: this sequence does not appear in the specification; it was created
CC using information provided.

XX Sequence 690 AA;

Query Match 93.3%; Score 56; DB 19; Length 690;

Best Local Similarity 91.7%; Pred. No. 0.08; Mismatches 1; Indels 0; Gaps 0;

Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ARRVRVWAAVG 12

|||||||

Db 339 arrarvvwcavg 350

RESULT 8

AAV58733

ID AAY58733 standard; Protein; 692 AA.
XX AAY58733;
AC AAY58733;
XX 25-APR-2000 (first entry)
DT 25-APR-2000 (first entry)
XX Human lactoferrin.
DE Human lactoferrin.
XX Lactoferrin; human; transgenic plant.
KW Lactoferrin; human; transgenic plant.
XX Homo sapiens.
OS Homo sapiens.
XX WO200004146-A1.
PN WO200004146-A1.
XX 27-JAN-2000.
PD 27-JAN-2000.
XX 19-JUL-1999; 99WO-IT00226.
PF 19-JUL-1999; 99WO-IT00226.
XX 17-JUL-1998; 98IT-RM00478.
PR 17-JUL-1998; 98IT-RM00478.
XX (PLAN-) PLANTECHNO SRL.
PA (PLAN-) PLANTECHNO SRL.
XX Fogher C;
PI Fogher C;
XX WPI; 2000-161129/14.
XX WPI; 2000-161129/14.
DR N-PSDB; AAZ58122.
DR N-PSDB; AAZ58122.
XX Synthetic polynucleotide encoding human lactoferrin, used for
PT production of functional foods, vegetal milks and human lactoferrin
PT production of functional foods, vegetal milks and human lactoferrin
XX Disclosure; Page 73-77; 93pp; English.
XX The present sequence is that of human lactoferrin. The invention
CC relates to a synthetic gene (see AAZ58122) that encodes human
CC lactoferrin but which has codon usage designed to maximise
CC expression in plants. Transgenic plants that express human
CC lactoferrin in a tissue-specific manner, especially in the seeds,
CC can be used in processes for the production of functional vegetal
CC milk, fruit juices, fruit and/or vegetable homogenized foods
CC (claimed). The transgenic plants are selected from solanaceae,
CC cereals, leguminosae, fruit bearing plants and horticultural
CC plants, especially soybean, tobacco and rice.
XX plants, especially soybean, tobacco and rice.
XX Sequence 692 AA;
SQ Sequence 692 AA;
Query Match 93.3%; Score 56; DB 21; Length 692;
Best Local Similarity 91.7%; Pred. No. 0.08;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ARRARVVAAG 12
DB 341 arrarvvwcavg 352
RESULT 9
AAB97382
ID AAB97382 standard; protein; 692 AA.
XX AAB97382;
AC AAB97382;
XX 17-AUG-2001 (first entry)
DT 17-AUG-2001 (first entry)
XX Human lactoferrin (hLF).
DE Human lactoferrin (hLF).
XX Human; lactoferrin; hLF; N-terminal; antimicrobial; heparin;
KW Human; lactoferrin; hLF; N-terminal; antimicrobial; heparin;
KW inflammatory response; cytokine production reduction;
KW neutrophil degranulation; myelopoiesis inhibition.
XX neutrophil degranulation; myelopoiesis inhibition.
OS Homo sapiens.
XX Homo sapiens.
XX Key Location/Qualifiers
FH Region 1..27
FT

FT /note= "Fragments of the N-terminal are specifically
FT referred to in the claims"
FT Domain 2.5
FT /label= Cationic_domain
FT Domain 28..31
FT /label= Cationic_domain
XX WO200134641-A2.
PN WO200134641-A2.
XX 17-MAY-2001.
PD 17-MAY-2001.
XX 10-NOV-2000; 2000WO-NL00821.
PF 10-NOV-2000; 2000WO-NL00821.
XX 11-NOV-1999; 99EP-0203775.
PR 11-NOV-1999; 99EP-0203775.
XX 11-NOV-1999; 99US-0164975.
PR 11-NOV-1999; 99US-0164975.
XX (PHAR-) PHARMING INTELLECTUAL PROPERTY BV.
PA (PHAR-) PHARMING INTELLECTUAL PROPERTY BV.
XX Van Berkel PHC, Nibbering PH, Nuijens JH;
PI Van Berkel PHC, Nibbering PH, Nuijens JH;
XX WPI; 2001-335909/35.
DR WPI; 2001-335909/35.
XX New polypeptides comprising the N-terminal region of human lactoferrin
PT protein exhibit higher antimicrobial activity than the full length
PT protein and are useful to treat bacterial infections
XX Claim 1; Page 55-57; 59pp; English.
PS Claim 1; Page 55-57; 59pp; English.
XX This invention relates to fragments (between 6 and 26 amino acids) of the
CC human lactoferrin hLF protein (represented by the present sequence).
CC human lactoferrin hLF protein (represented by the present sequence).
CC N-terminal hLF peptides have antimicrobial activity. The peptides of the
CC N-terminal hLF peptides have antimicrobial activity. The peptides of the
CC invention are used to treat microbial infections, especially infections
CC by gram positive or negative bacteria, particularly *Listeria*,
CC *Staphylococcus*, *Klebsiella* or *Escherichia* species, especially
CC *L. monocytogenes*, *S. aureus* and *E. coli*. Other uses include reducing
CC inflammatory response by neutralising heparin or lipopolysaccharide or by
CC reducing cytokine production and neutrophil degranulation, inhibiting
CC entry of viruses such as cytomegalovirus, herpes simplex virus 1 or HIV
CC into cells, inhibiting myelopoiesis and reducing production of GM-CSF
CC (granulocyte/macrophage colony stimulating factor).
XX (granulocyte/macrophage colony stimulating factor).
XX Sequence 692 AA;
SQ Sequence 692 AA;
Query Match 93.3%; Score 56; DB 22; Length 692;
Best Local Similarity 91.7%; Pred. No. 0.08;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ARRARVVAAG 12
DB 341 arrarvvwcavg 352
RESULT 10
AAR85146
ID AAR85146 standard; Protein; 693 AA.
XX AAR85146;
AC AAR85146;
XX 02-FEB-1996 (first entry)
DT 02-FEB-1996 (first entry)
XX Lactoferrin protein.
DE Lactoferrin protein.
XX Lactoferrin; antiseptic; dietary supplement; iron sequestration;
KW Lactoferrin; antiseptic; dietary supplement; iron sequestration;
KW *Pichia pastoris*.
XX *Pichia pastoris*.
OS Homo sapiens.
XX Homo sapiens.
XX WO9530339-A1.
PN WO9530339-A1.
XX 16-NOV-1995.
PD 16-NOV-1995.
XX 05-MAY-1995; 95WO-US05653.
FT 05-MAY-1995; 95WO-US05653.

XX PR 05-MAY-1994; 94US-0238445.
 XX (FERR-) FERRODYNAMICS INC.
 XX Doyle D, Gollnick P, Kruzel ML, Kurecki T;
 XX WPI; 1995-403881/51.
 XX N-PSDB; AAT02999.
 XX New isolated DNA encoding human lactoferrin protein - useful for
 PT inhibiting microbial growth and iron deficiency, and for
 PT sequestering iron or retarding food spoilage.
 XX Disclosure; Fig 3a-d; 43pp; English.
 XX cDNA encoding human lactoferrin (AAR85146) was isolated from a human
 CC mammary gland library and inserted into vector pUC118. Expression
 CC in Pichia pastoris GTS 115 (HIS4) allowed the prodn. of large
 CC amounts of the lactoferrin protein, useful as an antiseptic,
 CC dietary supplement, etc.
 XX Sequence 693 AA;

Query Match 93.3%; Score 56; DB 16; Length 693;
 Best Local Similarity 91.7%; Pred. No. 0.081; Mismatches 1; Indels 0; Gaps 0;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 AARARVWMAVG 12
 |||||
 DB 342 AARARVWMAVG 353

RESULT 11
 AAR22423
 ID AAR22423 standard; Protein; 705 AA.
 XX AAR22423;
 AC AAR22423;
 XX 05-AUG-1992 (first entry)
 XX Human lactoferrin deduced from clone HLF 1212.
 XX anti-bacterial; anti-viral; anti-fungal; Colony stimulating factor;
 KW CSP; release; leukaemia; breast cancer; hormonal regulation.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 FT Misc-difference 22 /note= "Insertion, c.f. B.F.Anderson et al, 1989"
 FT Misc-difference 31 /note= "Substitution, c.f. B.F.Anderson et al, 1989"
 FT Misc-difference 55 /note= "Substitution, c.f. B.F.Anderson et al, 1989"
 FT Misc-difference 218 /note= "Substitution, c.f. B.F.Anderson et al, 1989"
 XX US707502-A.
 PN 31-MAY-1991; 91US-0707502.
 PD 18-FEB-1992.
 XX 31-MAY-1991; 91US-0707502.
 XX 31-MAY-1991; 91US-0707502.
 XX (USSH) US DEPT HEALTH & HUMAN.
 PA Teng C, Panella TJ;
 XX WPI; 1992-113934/14.
 DR N-PSDB; AAR23453.

XX Human lactoferrin gene and promoter - used for producing protein
 PT for treating e.g. AIDS, neutropenia, skin infections, vaginal
 PT infections or septic shock
 XX Example 2; Fig 10; 55pp; English.
 XX A Clonotech cDNA library from normal human breast tissue (HL1037b)
 CC was plated in host cells 71090, filter-lifted and probed with mouse
 CC lactoferrin cDNA T267. Positive clones were plaque-purified and the
 CC inserts subcloned into the EcoRI site of Bluescript II SK+. The
 CC recombinant clones were transformed into XLI Blue cells. The
 CC longest insert (HLF 1212) was sequenced and was 2117bp in length.
 CC The amino acid sequence coded for by HLF 1212 has 4 sites that
 CC differ from the previously published revised amino acid sequence
 CC derived from the protein (B.F.Anderson et al., J.Mol.Biol. 209:
 CC 711-734, 1989). Also contained within HLF 1212, but not in any of
 CC the other partially sequenced isolates, is a deleted cytosine at
 CC bp 2097 (AA 699) which caused a frameshift at the 3' end of the coding
 CC sequence. The deletion is thought to be either a cloning artefact
 CC or a rare species of mRNA. See also AAR23454 and AAR23455.
 XX Sequence 705 AA;

Query Match 93.3%; Score 56; DB 13; Length 705;
 Best Local Similarity 91.7%; Pred. No. 0.082; Mismatches 1; Indels 0; Gaps 0;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 AARARVWMAVG 12
 |||||
 DB 358 AARARVWMAVG 369

RESULT 12
 AAY31152
 ID AAY31152 standard; Protein; 705 AA.
 XX AAY31152;
 AC AAY31152;
 XX 26-OCT-1999 (first entry)
 XX Human lactoferrin HLF1212 protein fragment.
 XX Lactoferrin; human; detection; primer; probe; malignant; cancer;
 KW screening; HLF1212; secretion.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 FT Protein 1.705
 FT /note= "Partial protein, no start or stop codon given"
 XX US5948613-A.
 PN 07-SEP-1999.
 PD 30-MAY-1996; 96US-0655640.
 XX 31-MAY-1991; 91US-0707502.
 PR 17-DEC-1992; 92US-0992538.
 PR 28-DEC-1994; 94US-0366006.
 PR 30-MAY-1996; 96US-0655640.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX Panella TJ, Teng C;
 XX WPI; 1999-517943/43.
 DR N-PSDB; AAZ09274.
 XX Screening for the risk of cancer using a DNA probe or primer
 XX

PS Example 2; Fig 10A-D; 3lpp; English.

CC This invention describes a novel method of screening for risk of cancer
 CC in a human by probing/priming test DNA with a human lactoferrin DNA
 CC probe/primer. The invention describes a human lactoferrin cDNA and its
 CC encoded protein obtained from human breast tissue. The products of the
 CC invention can be used in methods for detecting malignancy arising from
 CC tissues that normally secrete lactoferrin. This sequence represents the
 CC human lactoferrin designated HLF1212 described in the method of the
 CC invention.

XX Sequence 705 AA;

Query Match 93.3%; Score 56; DB 20; Length 705;

Best Local Similarity 91.7%; Pred. No. 0.082;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ARARVVWAAVG 12

Db 358 arrarvvwcavg 369
 ||||| |||

RESULT 13

AAR22424
 ID AAR22424 standard; Protein; 708 AA.

XX AAR22424;

DT 05-AUG-1992 (first entry)

DE Human lactoferrin deduced from clone HLF 1213.

XX anti-bacterial; anti-viral; anti-fungal; Colony stimulating factor;
 KW CSF; release; leukaemia; breast cancer; Hormonal regulation.

XX Homo sapiens.

XX US7707502-A.

XX 18-FEB-1992.

XX 31-MAY-1991; 91US-0707502.

PR 31-MAY-1991; 91US-0707502.

XX (USSH) US DEPT HEALTH & HUMAN.

XX Teng C, Panella TJ;

XX WPI; 1992-113934/14.

DR N-PSDB; AAQ23454.

XX Human lactoferrin gene and promoter - used for producing protein
 PT for treating e.g. AIDS, neutropenia, skin infections, vaginal
 PT infections or septic shock

PS Example 2; Page 32-36; 55pp; English.

XX A Clontech cDNA library from normal human breast tissue (HLF1037b)
 CC was plated in host cells 71090, filter-lifted and probed with mouse
 CC lactoferrin cDNA T267. Positive clones were plaque-purified and the
 CC inserts subcloned into the EcoRI site of Bluescript II SK+. The
 CC recombinant clones were transformed into XLI Blue cells. The
 CC longest insert (HLF 1212 - see AAQ23453) differed from HLF 1213 (and
 CC all the other partially sequenced isolates) in having a deleted
 CC cytosine which caused a frameshift at the 3' end of the coding
 CC sequence and consequently a change in amino acid sequence (c.f.
 CC AAR22423). The deletion is thought to be either a cloning artefact
 CC or a rare species of mRNA. See also AAQ23455.

XX Sequence 708 AA;

Query Match 93.3%; Score 56; DB 13; Length 708;
 Best Local Similarity 91.7%; Pred. No. 0.082;

Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ARARVVWAAVG 12

Db 360 arrarvvwcavg 371
 ||||| |||

RESULT 14

AAV31153
 ID AAV31153 standard; Protein; 708 AA.

XX AAV31153;

DT 26-OCT-1999 (first entry)

DE Human lactoferrin HLF1213 protein fragment.

XX Lactoferrin; human; detection; primer; probe; malignant; cancer;
 KW screening; HLF1213; secretion.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Protein 1..708

XX /note= "Partial sequence, no stop codon given"

XX US5948613-A.

XX 07-SEP-1999.

XX 30-MAY-1996; 96US-0655640.

PR 31-MAY-1991; 91US-0707502.

PR 17-DEC-1992; 92US-0992538.

PR 28-DEC-1994; 94US-0366006.

PR 30-MAY-1996; 96US-0655640.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Panella TJ, Teng C;

XX WPI; 1999-517943/43.

DR N-PSDB; AAQ09275.

XX Screening for the risk of cancer using a DNA probe or primer

XX Disclosure; Column 23-28; 3lpp; English.

XX This invention describes a novel method of screening for risk of cancer
 CC in a human by probing/priming test DNA with a human lactoferrin DNA
 CC probe/primer. The invention describes a human lactoferrin cDNA and its
 CC encoded protein obtained from human breast tissue. The products of the
 CC invention can be used in methods for detecting malignancy arising from
 CC tissues that normally secrete lactoferrin. This sequence represents a
 CC human lactoferrin designated HLF1213 described in the method of
 CC the invention.

XX Sequence 708 AA;

Query Match 93.3%; Score 56; DB 20; Length 708;

Best Local Similarity 91.7%; Pred. No. 0.082;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ARARVVWAAVG 12

Db 360 arrarvvwcavg 371
 ||||| |||

RESULT 15

AA012583
ID AA012583 standard; Protein; 709 AA.
XX
AC AA012583;
XX
DT 12-SEP-1991 (first entry)
XX
DE Lactoferrin.
XX
KW LF; baby food; milk; human.
XX
OS Homo sapiens.
XX
XX
XX Key Location/Qualifiers
XX Peptide 1..18
XX /label= partial signal sequence
XX Protein 19..709
XX /label= mature hLF
XX Misc-difference 140
XX /note= "Arg absent in ref 1"
XX Misc-difference 148
XX /note= "Thr - Ile in ref 1"
XX Misc-difference 169
XX /note= "Gln - Arg in ref 1"
XX Misc-difference 202
XX /note= "Ser - Leu in ref 1"
XX Misc-difference 207
XX /note= "Tyr - Lys in ref 1"
XX Misc-difference 390
XX /note= "Ser - Trp in ref 1"
XX Misc-difference 405..406
XX /note= "Insertion of 13 AAs in ref 1"
XX Misc-difference 421
XX /note= "Cys - Gly in ref 1"
XX Misc-difference 530
XX /note= "Gln - Glu in ref 1"
XX Misc-difference 693
XX /note= "Lys - Arg in ref 1"
XX
XX
XX W09108216-A.
XX
XX 13-JUN-1991.
XX
XX 30-NOV-1990; 90WO-US06874.
XX
XX 01-DEC-1989; 89US-0444745.
XX
XX (GENP-) GENPHARM INT.
XX
XX Heyneker HL, Deboer HA, Strijker R, Platenbur G, Lee SH;
XX
XX WPI; 1991-193146/26.
XX
XX N-PSDB; AAQ12230.
XX
XX Trans-gene for prodn. of recombinant polypeptide(s) in bovine
XX milk - prodn. of transgenic animals with desired phenotype, and
XX use of recombinant polypeptide(s) in food.
XX
XX Disclosure; Fig 1; 120pp; English.
XX
XX DNA encoding the protein can be ligated with regulatory and secret-
XX ory sequences to prepare an expression vector which can be used to
XX produce a transgenic bovine species. The recombinant lactoferrin is
XX expressed in bovine milk from which it can be purified and used to
XX supplement e.g. baby food. A no. of discrepancies exist between
XX this sequence and that disclosed by Metz-Boutigue et al (1984), Eur.
XX J. Biochem., 21451, 659-676).
XX
XX Sequence 709 AA;
XX

Query Match 93.3%; Score 56; DB 12; Length 709;
Best Local Similarity 91.7%; Pred. No. 0.082;

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 31, 2001, 08:52:06 ; Search time 12.67 seconds
(without alignments)
21.313 Million cell updates/sec

Title: US-09-508-095-22
Perfect score: 60
Sequence: 1 ARRVRVVAAGV 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2.6/ptodata/2/1aa/5A_COMB.pep.*
2: /cgn2.6/ptodata/2/1aa/5B_COMB.pep.*
3: /cgn2.6/ptodata/2/1aa/6A_COMB.pep.*
4: /cgn2.6/ptodata/2/1aa/6B_COMB.pep.*
5: /cgn2.6/ptodata/2/1aa/PCTUS_COMB.pep.*
6: /cgn2.6/ptodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	56	93.3	694	US-08-724-586-2
2	56	93.3	694	US-09-421-632-2
3	56	93.3	705	US-08-655-640-2
4	56	93.3	708	US-08-655-640-4
5	56	93.3	709	US-08-154-019-2
6	56	93.3	709	US-08-461-333-2
7	56	93.3	709	US-08-464-167-2
8	56	93.3	709	US-09-158-313-2
9	56	93.3	709	US-08-476-798-2
10	56	93.3	711	US-08-145-681-2
11	56	93.3	711	US-08-250-308-2
12	56	93.3	711	US-08-154-019-4
13	56	93.3	711	US-08-461-333-4
14	56	93.3	711	US-08-453-703-2
15	56	93.3	711	US-08-456-106-2
16	56	93.3	711	US-08-464-167-4
17	56	93.3	711	US-09-158-313-4
18	56	93.3	711	US-08-456-108-2
19	56	93.3	711	US-08-476-798-4
20	56	93.3	711	US-09-265-577-2
21	56	93.3	711	PCT-US93-03614-2
22	49	81.7	703	US-08-145-681-6
23	49	81.7	703	US-08-453-703-6
24	49	81.7	703	US-08-456-106-6
25	49	81.7	703	US-08-456-108-6
26	49	81.7	703	US-09-265-577-6
27	45	75.0	708	US-08-145-681-4

28 45 75.0 708 1 US-08-453-703-4 Sequence 4, Appli
29 45 75.0 708 2 US-08-456-106-4 Sequence 4, Appli
30 45 75.0 708 3 US-08-456-108-4 Sequence 4, Appli
31 45 75.0 708 4 US-09-265-577-4 Sequence 4, Appli
32 40 66.7 328 2 US-08-977-767-1 Sequence 1, Appli
33 38 63.3 284 1 US-08-061-889-2 Sequence 2, Appli
34 38 63.3 284 1 US-08-462-611-2 Sequence 2, Appli
35 38 63.3 284 5 PCT-US94-05378-2 Sequence 2, Appli
36 37 61.7 427 4 US-09-199-737-4 Sequence 4, Appli
37 35 58.3 20 1 US-07-678-974D-9 Sequence 9, Appli
38 35 58.3 20 2 US-08-945-168-14 Sequence 14, Appli
39 35 58.3 505 4 US-08-944-368A-2 Sequence 2, Appli
40 34 56.7 38 1 US-08-318-193-81 Sequence 81, Appli
41 33 55.0 220 2 US-08-726-306A-29 Sequence 29, Appli
42 33 55.0 303 4 US-09-372-934-2 Sequence 2, Appli
43 33 55.0 339 3 US-08-968-563-14 Sequence 14, Appli
44 33 55.0 339 4 US-08-969-683A-14 Sequence 14, Appli
45 33 55.0 903 1 US-08-220-151-8 Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-08-724-586-2
; Sequence 2, Application US/08724586
; Patent No. 6066469
; GENERAL INFORMATION:
; APPLICANT: Kruzel, Marian L.
; APPLICANT: Kurecki, Tomasz
; APPLICANT: Gollnick, Paul D.
; APPLICANT: Doyle, Darrell J.
; TITLE OF INVENTION: Cloning, Expression, and Uses of Human
; TITLE OF INVENTION: Lactoferrin
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jacobson, Price, Holman & Stern
; STREET: 400 Seventh St. N.W.
; CITY: Washington D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/724,586
; APPLICATION NUMBER: US/08/724,586
; FILING DATE: 30-SEPT-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/238,445
; FILING DATE: 05-MAY-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Player, William E.
; REGISTRATION NUMBER: 31,409
; REFERENCE/DOCKET NUMBER: 10505/P58185C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 638-6666
; TELEFAX: (202) 393-5350
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 694 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-724-586-2

Query Match 93.3%; Score 56; DB 3; Length 694;
Best Local Similarity 91.7%; Pred. No. 0.055;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ARRVRVWVAAG 12
Db 343 ARRVRVWCAVG 354

RESULT 2
US-09-421-632-2
; Sequence 2, Application US/09421632
; Patent No. 6277817
; GENERAL INFORMATION:
; APPLICANT: Kruzel, Marian L.
; APPLICANT: Kurecki, Tomasz
; APPLICANT: Gollnick, Paul D.
; APPLICANT: Doyle, Darrell J.
; TITLE OF INVENTION: Cloning, Expression, and Uses of Human
; TITLE OF INVENTION: Lactoferrin
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jacobson, Price, Holman & Stern
; STREET: 400 Seventh St. N.W.
; CITY: Washington D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/421.632
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/724,586
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Player, William E.
; REGISTRATION NUMBER: 31,409
; REFERENCE/DOCKET NUMBER: 10505/P58185C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 638-6666
; TELEFAX: (202) 393-5350
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 694 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-421-632-2

Query Match 93.3%; Score 56; DB 4; Length 694;
Best Local Similarity 91.7%; Pred. No. 0.055;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ARRVRVWVAAG 12
Db 343 ARRVRVWCAVG 354

RESULT 3
US-08-655-640-2
; Sequence 2, Application US/08655640
; Patent No. 5948613
; GENERAL INFORMATION:
; APPLICANT: Teng, Christina
; APPLICANT: Panella, Timothy J.
; TITLE OF INVENTION: HUMAN LACTOFERRIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
; STREET: 1100 NEW YORK AVE. N.W., NINTH FLOOR
; CITY: WASHINGTON

STATE: D.C.
COUNTRY: USA
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/655,640
FILING DATE: 30-MAY-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/992,538
FILING DATE: December 17, 1992
ATTORNEY/AGENT INFORMATION:
NAME: SCOTT, WATSON T.
REGISTRATION NUMBER: 26,581
REFERENCE/DOCKET NUMBER: WTS/5683/98019/SAP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 861-3000
TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 705 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-655-640-2

Query Match 93.3%; Score 56; DB 2; Length 705;
Best Local Similarity 91.7%; Pred. No. 0.056;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ARRVRVWVAAG 12
Db 358 ARRVRVWCAVG 369

RESULT 4
US-08-655-640-4
; Sequence 4, Application US/08655640
; Patent No. 5948613
; GENERAL INFORMATION:
; APPLICANT: Teng, Christina
; APPLICANT: Panella, Timothy J.
; TITLE OF INVENTION: HUMAN LACTOFERRIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
; STREET: 1100 NEW YORK AVE. N.W., NINTH FLOOR
; CITY: WASHINGTON
; STATE: D.C.
COUNTRY: USA
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/655,640
FILING DATE: 30-MAY-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/992,538
FILING DATE: December 17, 1992
ATTORNEY/AGENT INFORMATION:
NAME: SCOTT, WATSON T.
REGISTRATION NUMBER: 26,581
REFERENCE/DOCKET NUMBER: WTS/5683/98019/SAP

TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 861-3000
TELEFAX: (202) 822-0944
TELEX: 6710627 CUSH
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 708 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-655-640-4

Query Match 93.3%; Score 56; DB 2; Length 708;
Best Local Similarity 91.7%; Pred. No. 0.056;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ARRVRVVAAG 12
| | | | | | | | | | | | | |
Db 360 ARRVRVVAAG 371

RESULT 5
US-08-154-019-2
Sequence 2, Application US/08154019
Patent No. 5633076
GENERAL INFORMATION:
APPLICANT: Deboer, Herman A.
APPLICANT: Strijker, Rein
APPLICANT: Heyneker, Herbert L.
APPLICANT: Platenburg, Gerald
APPLICANT: Lee, Sang He
APPLICANT: Pieper, Frank
APPLICANT: Krimpenfort, Paul J.A.
TITLE OF INVENTION: Production of Recombinant Polypeptides
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew
STREET: One Market Plaza, Steuart Tower, Suite 2000
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/154,019
FILING DATE: 16-NOV-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/461,333
FILING DATE: 05-JUN-1995
APPLICATION NUMBER: US 08/077,788
FILING DATE: 15-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/895,956
FILING DATE: 15-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/619,131
FILING DATE: 27-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/444,745
FILING DATE: 01-DEC-1989
ATTORNEY/AGENT INFORMATION:
NAME: Liebescheutz, Joe O.
REGISTRATION NUMBER: 37,505
REFERENCE/DOCKET NUMBER: 16994-003123
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-543-9600

TELEFAX: 415-543-5043
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 709 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-154-019-2
Query Match 93.3%; Score 56; DB 1; Length 709;
Best Local Similarity 91.7%; Pred. No. 0.056;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ARRVRVVAAG 12
| | | | | | | | | | | | | |
Db 358 ARRVRVVAAG 369
RESULT 6
US-08-461-333-2
Sequence 2, Application US/08461333
Patent No. 5741957
GENERAL INFORMATION:
APPLICANT: Deboer, Herman A.
APPLICANT: Strijker, Rein
APPLICANT: Heyneker, Herbert L.
APPLICANT: Platenburg, Gerald
APPLICANT: Lee, Sang He
APPLICANT: Pieper, Frank
APPLICANT: Krimpenfort, Paul J.A.
TITLE OF INVENTION: Production of Recombinant Polypeptides
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew
STREET: One Market Plaza, Steuart Tower, Suite 2000
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/461,333
FILING DATE: 05-JUN-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/077,788
FILING DATE: 15-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/895,956
FILING DATE: 15-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/619,131
FILING DATE: 27-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/444,745
FILING DATE: 01-DEC-1989
ATTORNEY/AGENT INFORMATION:
NAME: Liebescheutz, Joe O.
REGISTRATION NUMBER: 37,505
REFERENCE/DOCKET NUMBER: 16994-003123
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-543-9600
TELEFAX: 415-543-5043
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 709 amino acids
TYPE: amino acid

; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-461-333-2

Query Match 93.3%; Score 56; DB 1; Length 709;
Best Local Similarity 91.7%; Pred. No. 0.056;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ARRVRVWAAVG 12
|||||

Db 358 ARRVRVWCAVG 369

RESULT 7

US-08-464-167-2
; Sequence 2, Application US/08464167
; Patent No. 6013857
; GENERAL INFORMATION:
; APPLICANT: Deboer, Herman A.
; APPLICANT: Strijker, Rein
; APPLICANT: Heyneker, Herbert L.
; APPLICANT: Platenburg, Gerald
; APPLICANT: Lee, Sang He
; APPLICANT: Pieper, Frank
; APPLICANT: Krimpenfort, Paul J.A.
; TITLE OF INVENTION: Production of Recombinant Polypeptides
; TITLE OF INVENTION: by Bovine Species and Transgenic Methods
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew
; STREET: One Market Plaza, Steuart Tower, Suite 2000
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/464,167
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/077,788
; FILING DATE: 15-JUN-1993

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/895,956
; FILING DATE: 15-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/619,131
; FILING DATE: 27-NOV-1990

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/444,745
; FILING DATE: 01-DEC-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Liebescheutz, Joe O.
; REGISTRATION NUMBER: 37,505
; REFERENCE/DOCKET NUMBER: 16994-003124
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-543-9600
; TELEFAX: 415-543-5043

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 709 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

; MOLECULE TYPE: protein
US-08-464-167-2

Query Match 93.3%; Score 56; DB 3; Length 709;
Best Local Similarity 91.7%; Pred. No. 0.056;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ARRVRVWAAVG 12
|||||

Db 358 ARRVRVWCAVG 369

RESULT 8

US-09-158-313-2
; Sequence 2, Application US/09158313
; Patent No. 6086725
; GENERAL INFORMATION:
; APPLICANT: Deboer, Herman A.
; APPLICANT: Strijker, Rein
; APPLICANT: Heyneker, Herbert L.
; APPLICANT: Platenburg, Gerald
; APPLICANT: Lee, Sang He
; APPLICANT: Pieper, Frank
; APPLICANT: Krimpenfort, Paul J.A.
; TITLE OF INVENTION: Production of Recombinant Polypeptides
; TITLE OF INVENTION: by Bovine Species and Transgenic Methods
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew
; STREET: One Market Plaza, Steuart Tower, Suite 2000
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/09/158,313
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/476,798
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/895,956
; FILING DATE: 15-JUN-1992

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/619,131
; FILING DATE: 27-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/444,745
; FILING DATE: 01-DEC-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Liebescheutz, Joe O.
; REGISTRATION NUMBER: 37,505
; REFERENCE/DOCKET NUMBER: 16994-003125
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-543-9600
; TELEFAX: 415-543-5043

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 709 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-158-313-2

Query Match 93.3%; Score 56; DB 3; Length 709;
Best Local Similarity 91.7%; Pred. No. 0.056;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ARRVRVWAAVG 12

Db 358 ARRARVWCAVG 369
|||||

RESULT 9
 US-08-476-798-2
 ; Sequence 2, Application US/08476798
 ; Patent No. 6140552
 ;
 ; GENERAL INFORMATION:
 ; APPLICANT: Deboer, Herman A.
 ; APPLICANT: Strijker, Rein
 ; APPLICANT: Heyneker, Herbert L.
 ; APPLICANT: Platenburg, Gerald
 ; APPLICANT: Lee, Sang He
 ; APPLICANT: Pieper, Frank
 ; APPLICANT: Krimpenfort, Paul J.A.
 ; TITLE OF INVENTION: Production of recombinant Polypeptides
 ; ; by Bovine Species and Transgenic Methods
 ; NUMBER OF SEQUENCES: 38
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Townsend and Townsend and Crew
 ; STREET: One Market Plaza, Steuart Tower, Suite 2000
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94105
 ;
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ;
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/476.798
 ; FILING DATE: 07-JUN-1995
 ; CLASSIFICATION: 800
 ;
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/077,788
 ; FILING DATE: 15-JUN-1993
 ;
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/895,956
 ; FILING DATE: 15-JUN-1992
 ;
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/619,131
 ; FILING DATE: 27-NOV-1990
 ;
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/444,745
 ; FILING DATE: 01-DEC-1989
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Liebeschultz, Joe O.
 ; REGISTRATION NUMBER: 37,505
 ; REFERENCE/DOCKET NUMBER: 16994-003125
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415-543-9600
 ; TELEFAX: 415-543-5043
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 709 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 US-08-476-798-2

```

US-08-145-681-2
; Sequence 2, Application US/08145681
; Patent No. 5571691
; GENERAL INFORMATION:
; APPLICANT: Conneely, Orla M.
; APPLICANT: Headon, Denis R.
; APPLICANT: O'Malley, Bert W.
; APPLICANT: May, Gregory S.
; TITLE OF INVENTION: Production of Recombinant Lactoferrin
; and Lactoferrin Polypeptides Using CDNA Sequences in
; TITLE OF INVENTION: Various Organsims
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Baker & Botts, L.L.P.
; STREET: 910 Louisiana St
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77002
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/145,681
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McGregor, Martin L.
; REGISTRATION NUMBER: 29,329
; REFERENCE/DOCKET NUMBER: 19928-0125
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 713/229/1874
; TELEFAX: 713/229/1522
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 711 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: H. sapiens
US-08-145-681-2

```

```

Query Match      93.3%; Score 56; DB 4; Length 709;
Best Local Similarity 91.7%; Pred. No. 0.056;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1  AARRRVVWAAYG 12
          |||||
Db       358  AARRRVVWCAVG 369

RESULT 10

```

STATE: Texas
COUNTRY: USA
ZIP: 77010-3095
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/250,308
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/873,304
FILING DATE: 24-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Adler Ph.D., Benjamin A.
REGISTRATION NUMBER: 35,423
REFERENCE/DOCKET NUMBER: D5456
TELECOMMUNICATION INFORMATION:
TELEPHONE: 7136515587
TELEFAX: 7136515246
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 711 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-250-308-2

Query Match 93.3%; Score 56; DB 1; Length 711;
Best Local Similarity 91.7%; Pred. No. 0.056;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ARRVRVVAAG 12
Db 360 ARRVRVVAAG 371
|||||

RESULT 12
US-08-154-019-4
Sequence 4, Application US/08154019
Patent No. 5633076
GENERAL INFORMATION:
APPLICANT: Deboer, Herman A.
APPLICANT: Strijker, Rein
APPLICANT: Heyneker, Herbert L.
APPLICANT: Platenburg, Gerald
APPLICANT: Lee, Sang He
APPLICANT: Pieper, Frank
APPLICANT: Krimpenfort, Paul J.A.
TITLE OF INVENTION: Production of Recombinant Polypeptides
TITLE OF INVENTION: by Bovine Species and Transgenic Methods
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Crew
STREET: One Market Plaza, Steuart Tower, Suite 2000
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/154,019
FILING DATE: 16-NOV-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/461,333

FILING DATE: 05-JUN-1995
APPLICATION NUMBER: US 08/077,788
FILING DATE: 15-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/895,956
FILING DATE: 15-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/619,131
FILING DATE: 27-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/444,745
FILING DATE: 01-DEC-1989
ATTORNEY/AGENT INFORMATION:
NAME: Liebescheutz, Joe O.
REGISTRATION NUMBER: 37,505
REFERENCE/DOCKET NUMBER: 16994-003123
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-543-9600
TELEFAX: 415-543-5043
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 711 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-154-019-4

Query Match 93.3%; Score 56; DB 1; Length 711;
Best Local Similarity 91.7%; Pred. No. 0.056;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ARRVRVVAAG 12
Db 360 ARRVRVVAAG 371
|||||

RESULT 13
US-08-461-333-4
Sequence 4, Application US/08461333
Patent No. 5741957
GENERAL INFORMATION:
APPLICANT: Deboer, Herman A.
APPLICANT: Strijker, Rein
APPLICANT: Heyneker, Herbert L.
APPLICANT: Platenburg, Gerald
APPLICANT: Lee, Sang He
APPLICANT: Pieper, Frank
APPLICANT: Krimpenfort, Paul J.A.
TITLE OF INVENTION: Production of Recombinant Polypeptides
TITLE OF INVENTION: by Bovine Species and Transgenic Methods
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Crew
STREET: One Market Plaza, Steuart Tower, Suite 2000
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/461,333
FILING DATE: 05-JUN-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/077,788
FILING DATE: 15-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/895,956

; FILING DATE: 15-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/619,131
; FILING DATE: 27-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/444,745
; FILING DATE: 01-DEC-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Liebescheutz, Joe O.
; REGISTRATION NUMBER: 37,505
; REFERENCE/DOCKET NUMBER: 16994-003123
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-543-9600
; TELEFAX: 415-543-5043
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 711 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-461-333-4

Query Match 93.3%; Score 56; DB 1; Length 711;
Best Local Similarity 91.7%; Pred. No. 0.056;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 ARRARVWVAAG 12
Db 360 ARRARVWCAVG 371

RESULT 14
US-08-453-703-2
; Sequence 2, Application US/08453703
; Patent No. 5766939
; GENERAL INFORMATION:
; APPLICANT: Conneely, Orla M.
; APPLICANT: Headon, Denis R.
; APPLICANT: O'Malley, Bert W.
; APPLICANT: May, Gregory S.
; TITLE OF INVENTION: Production of Recombinant Lactoferrin
; TITLE OF INVENTION: and Lactoferrin Polypeptides Using cdna Sequences in
; TITLE OF INVENTION: Various Organisms
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/453,703
; FILING DATE: Concurrently herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/145,681
; FILING DATE: October 28, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Albert P. Halluin
; REGISTRATION NUMBER: 25,227
; REFERENCE/DOCKET NUMBER: 8206-024
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-854-3660
; TELEFAX: 415-854-3694
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 711 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: H. sapiens
US-08-453-703-2

; LENGTH: 711 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: H. sapiens
US-08-453-703-2

Query Match 93.3%; Score 56; DB 1; Length 711;
Best Local Similarity 91.7%; Pred. No. 0.056;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 ARRARVWVAAG 12
Db 360 ARRARVWCAVG 371

RESULT 15
US-08-456-106-2
; Sequence 2, Application US/08456106
; Patent No. 5849881
; GENERAL INFORMATION:
; APPLICANT: Conneely, Orla M.
; APPLICANT: Headon, Denis R.
; APPLICANT: O'Malley, Bert W.
; APPLICANT: May, Gregory S.
; TITLE OF INVENTION: Production of Recombinant Lactoferrin
; TITLE OF INVENTION: and Lactoferrin Polypeptides Using cdna Sequences in
; TITLE OF INVENTION: Various Organisms
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/456,106
; FILING DATE: Concurrently herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/145,681
; FILING DATE: October 28, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Albert P. Halluin
; REGISTRATION NUMBER: 25,227
; REFERENCE/DOCKET NUMBER: 8206-025
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-854-3660
; TELEFAX: 415-854-3694
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 711 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: H. sapiens
US-08-456-106-2

Query Match 93.3%; Score 56; DB 2; Length 711;

Best Local Similarity 91.7%; Pred. No. 0.056;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ARRARVVWAAVG 12
| | | | | | | | | |
Db 360 ARRARVVWCAVG 371

Search completed: December 31, 2001, 08:52:29
Job time: 23 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 31, 2001, 08:52:06 ; Search time 12.88 Seconds
(without alignments)
70.970 Million cell updates/sec

Title: US-09-508-095-22
Perfect score: 60
Sequence: 1 ARRARNVWAAG 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	56	93.3	711	1 TFHUL	lactotransferrin p
2	49	81.7	703	2 A45543	lactoferrin precu
3	45	75.0	708	1 TFBOL	lactotransferrin p
4	44	73.3	708	2 JC2323	lactoferrin - goat
5	43	71.7	491	2 T46684	p-hydroxybenzaldehy
6	43	71.7	707	1 A28438	lactoferrin precu
7	40	66.7	328	2 JE0375	carbonic anhydrase
8	40	66.7	449	1 H70526	probable cytochrom
9	39	65.0	285	1 IDECRP	replication initia
10	39	65.0	285	2 I64780	replication-associ
11	38	63.3	176	2 F72563	hypothetical prote
12	38	63.3	509	2 T02260	beta-fructofuranos
13	37	61.7	108	2 F72507	hypothetical prote
14	37	61.7	338	2 T12816	hypothetical prote
15	37	61.7	397	2 A75803	hypothetical prote
16	37	61.7	419	1 S74500	zeaxanthin glucosy
17	37	61.7	477	2 D83617	probable amino aci
18	37	61.7	651	2 A26581	beta-glucuronidase
19	36	60.0	148	2 F86615	cytosine deaminase
20	36	60.0	155	2 E72007	cytosine deaminase
21	36	60.0	342	2 S76826	ABC-type transport
22	36	60.0	399	2 T30222	sensory protein ki
23	36	60.0	497	2 T48676	proline uptake pro
24	36	60.0	501	1 P1WLEP	L1 protein - Europ
25	36	60.0	507	2 S36478	L1 protein - human
26	36	60.0	513	1 P1WLDP	L1 protein - deer
27	36	60.0	771	1 W2VZ84	ribonucleoside-dip
28	36	60.0	771	1 WMV29J	ribonucleoside-dip
29	36	60.0	771	2 T28496	ribonucleoside-dip

ALIGNMENTS

RESULT 1

TFHUL
lactotransferrin precursor [validated] - human

N;Alternate names: lactoferrin

C;Species: Homo sapiens (man)

C;Date: 31-Mar-1992 #sequence_revision 21-Nov-1997 #text_change 08-Dec-2000

C;Accession: G01394; S11228; A45401; S10324; S15853; S20841; S07160; A61169; A31000;

R;Cho, Y.

submitted to the EMBL Data Library, March 1994

A;Reference number: G06820

A;Accession: G01394

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-711 <CHO>

A;Cross-references: EMBL:U07643; NID:g467236; PIDN:AAB60324.1; PID:g467237

R;Rey, M.W.; Woloshuk, S.L.; deBoer, H.A.; Pieper, F.R.

Nucleic Acids Res. 18, 5288, 1990

A;Title: Complete nucleotide sequence of human mammary gland lactoferrin.

A;Reference number: S11228; MUID:90384839

A;Accession: S11228

A;Molecule type: mRNA

A;Residues: 1-148,'T',150-422,'C',424-711 <REY>

A;Cross-references: EMBL:X53961; NID:g34415; PIDN:CAA37914.1; PID:g34416

R;Teng, C.T.; Liu, Y.; Yang, N.; Walmer, D.; Panella, T.

Mol. Endocrinol. 6, 1969-1981, 1992

A;Title: Differential molecular mechanism of the estrogen action that regulates lacto

A;Reference number: A45401; MUID:93125571

A;Accession: A45401

A;Molecule type: DNA

A;Residues: 1-15 <TEN>

A;Cross-references: GB:S52659; NID:g263311; PIDN:AAB24877.1; PID:g263312

A;Experimental source: placenta

A;Note: sequence extracted from NCBI backbone (NCBIP:122202).

R;Powell, M.J.; Ogden, J.E.

Nucleic Acids Res. 18, 4013, 1990

A;Title: Nucleotide sequence of human lactoferrin cDNA.

A;Reference number: S10324; MUID:90326549

A;Accession: S10324

A;Molecule type: mRNA

A;Residues: 3-711 <POW>

A;Cross-references: EMBL:X52941; NID:g34411; PIDN:CAA37116.1; PID:g34412

R;Stowell, K.M.; Rado, T.A.; Funk, W.D.; Tweedie, J.W.

Biochem. J. 276, 349-355, 1991

A;Title: Expression of cloned human lactoferrin in baby-hamster kidney cells.

A;Reference number: S15853; MUID:91264786

A;Accession: S15853

A;Status: nucleic acid sequence not shown; not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 20-31 <ST1>

A;Accession: S20841

A;Molecule type: protein

A;Residues: 20-28,'X',30-31 <ST2>

ribonucleoside-dip
L4L protein - vari
probable transcrip
type IIS restricti
ribonucleoside-dip
hypothetical prote
Mutator-like trans
probable integrase
probable metallopro
L1 protein - bovin
L1 protein - human
L1 protein - human
L1 protein - human
L1 protein - human
L1 protein - human

30 36 60.0 771 2 B36843
31 36 60.0 771 2 H72157
32 36 60.0 827 2 H83217
33 36 60.0 1279 2 E64709
34 35 58.3 57 2 C41476
35 35 58.3 84 2 D70967
36 35 58.3 145 2 C84498
37 35 58.3 402 2 T03567
38 35 58.3 457 2 D82961
39 35 58.3 495 1 P1WLB
40 35 58.3 499 1 P1WL33
41 35 58.3 499 2 S36531
42 35 58.3 501 1 P1WLB2
43 35 58.3 502 1 P1WL42
44 35 58.3 503 2 S36514
45 35 58.3 503 2 S36549

R;Rado, T.A.; Wei, X.; Benz Jr., E.J.
 Blood 70, 989-993, 1987
 A:Title: Isolation of lactoferrin cDNA from a human myeloid library and expression of mRNA
 A:Reference number: S07160; MUID:88001031
 A:Accession: S07160
 A:Molecule type: mRNA
 A:Residues: 436-487, 'A', 489-711 <RND>
 A:Cross-references: EMBL:M18642; NID:g186815; PIDN:AAA86665.1; PID:g386855
 R;Panella, T.J.; Liu, Y.; Huang, A.T.; Teng, C.T.
 Cancer Res. 51, 3037-3043, 1991
 A:Title: Polymorphism and altered methylation of the lactoferrin gene in normal leukocytes
 A:Reference number: A61169; MUID:91235214
 A:Accession: A61169
 A:Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 3-701, 'SWKPNV' <PAN>
 A:Experimental source: normal breast tissue
 R;Metz-Boutigue, M.H.; Jolles, J.; Mazurier, J.; Schoentgen, F.; Legrand, D.; Spik, G.;
 Eur. J. Biochem. 145, 659-666, 1984
 A:Title: Human lactotransferrin: amino acid sequence and structural comparisons with other
 A:Reference number: A31000; MUID:85076667
 A:Accession: A31000
 A:Molecule type: protein
 A:Residues: 20-140, 142-169, 171-203, 'L', 205, 'K', 207-208, 'K', 210-385, 'Q', 387-391, 'W', 393-4
 A:Note: this is the final paper in a series
 R;Houen, G.; Hoegdall, E.V.; Barkholt, V.; Norskov, L.
 Eur. J. Biochem. 241, 303-308, 1996
 A:Title: Lactoferrin: similarity to diamine oxidase and purification by aminohexyl affinity
 A:Reference number: S74119; MUID:97054624
 A:Accession: S74119
 A:Molecule type: protein
 A:Residues: 'G', 23-24, 'R', 26-27, 'XX', 30-32 <HOU>
 A:Experimental source: neutrophil granulocytes
 C:Genetics:
 A:Gene: GDB:ITF
 A:Cross-references: GDB:l19368; OMIM:150210
 A:Map position: 3q21-3q23
 C:Superfamily: transferrin; transferrin repeat homology
 C:Keywords: duplication; glycoprotein; iron binding; milk
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-71/Product: lactotransferrin #status experimental <MAT>
 F:21-356/Domain: transferrin repeat homology <TRH1>
 F:360-699/Domain: transferrin repeat homology <TRH2>
 F:29-65, 39-56, 135-218, 177-193, 190-201, 251-265, 503-697, 595-609/Disulfide bonds: #status e
 F:157, 498/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F:368-400, 378-391, 425-706, 447-669, 479-554, 513-527, 524-537, 647-652/Disulfide bonds: #stat

Query Match 93.3%; Score 56; DB 1; Length 711;
 Best Local Similarity 91.7%; Pred. No. 0.025;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ARARVVAWVG 12
 DB 360 ARARVVAWVG 371
 |||||
 RESULT 2
 Lactoferrin precursor - pig
 C:Species: Sus scrofa domestica (domestic pig)
 C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 04-Mar-2000
 C:Accession: A45543; S24173
 R;Alexander, L.J.; Levine, W.B.; Teng, C.T.; Beattie, C.W.
 Anim. Genet. 23, 251-256, 1992
 A:Title: Cloning and sequencing of the porcine lactoferrin cDNA.
 A:Reference number: A45543; MUID:92367939
 A:Accession: A45543
 A:Status: Preliminary
 A:Molecule type: mRNA
 A:Residues: 1-703 <ALE>
 A:Experimental source: mammary gland
 A:Note: sequence extracted from NCBI backbone (NCBIN:l11151, NCBI:P:l11153)

R;Lydon, J.P.; O'Malley, B.R.; Saucedo, O.; Lee, T.; Headon, D.R.; Conneely, O.M.
 Biochim. Biophys. Acta 1132, 97-99, 1992
 A:Title: Nucleotide and primary amino acid sequence of porcine lactoferrin.
 A:Reference number: S24173; MUID:92379101
 A:Accession: S24173
 A:Molecule type: mRNA
 A:Residues: 1-11, 'W', 13-50, 'I', 52-84, 'G', 86-120, 'L', 121-130, 'I', 132-282, 'S', 284-571, '
 A:Cross-references: EMBL:M92089; NID:g164613; PIDN:AAA31102.1; PID:g164614
 A:Experimental source: mammary gland
 C:Superfamily: transferrin; transferrin repeat homology
 C:Keywords: duplication; glycoprotein; iron; iron binding; metalloprotein; milk
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-703/Product: lactoferrin #status predicted <MAT>
 F:20-350/Domain: transferrin repeat homology <TRH1>
 F:36-48/Region: antimicrobial
 F:354-691/Domain: transferrin repeat homology <TRH2>
 F:28-62, 38-53, 129-212, 171-187, 184-195, 245-259, 362-394, 372-385, 419-698, 439-661, 471-546
 F:77, 107, 206, 267/Binding site: iron (Asp, Tyr, His) #status predicted
 F:335/Binding site: carbonate (Arg) #status predicted
 F:409, 447, 540, 609/Binding site: iron (Asp, Tyr, His) #status predicted
 F:477/Binding site: carbonate (Arg) #status predicted
 F:490/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 81.7%; Score 49; DB 2; Length 703;
 Best Local Similarity 75.0%; Pred. No. 0.45;
 Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ARARVVAWVG 12
 DB 354 ARARVVAWVG 365
 |||||
 RESULT 3
 TFBOL
 Lactotransferrin precursor - bovine
 N:Alternate names: lactoferrin
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 31-Mar-1992 #sequence_revision 21-Nov-1997 #text_change 11-May-2000
 C:Accession: I45919; S14674; S14110; S18517; JTO595; S13097; S18518; S13881; PL0148;
 R;Tsang, T.C.; Burns, D.K.; Wang, F.; Pan, Y.
 FASEB J. 6, 233, 1991
 A:Title: Cloning of a 80-kD advanced glycosylation end product (AGE) binding protein
 A:Reference number: I45919
 A:Accession: I45919
 A:Status: translated from GB/EMBL/DDBJ
 A:Molecule type: mRNA
 A:Residues: 1-708 <TSA>
 A:Cross-references: GB:L08604; NID:g163269; PIDN:AAA30609.1; PID:g163270
 R;Pierce, A.
 submitted to the EMBL Data Library, November 1990
 A:Reference number: S14674
 A:Accession: S14674
 A:Molecule type: mRNA
 A:Residues: 1-144, 'V', 146-163, 'PP', 166-339, 'A', 341-438, 'Y', 440-513, 'R', 515-708 <PI1>
 A:Cross-references: EMBL:X57084; NID:g505; PIDN:CAA40366.1; PID:g506
 R;Pierce, A.; Colavizza, D.; Benalissa, M.; Maes, P.; Tartar, A.; Montreuil, J.; Spik,
 Eur. J. Biochem. 196, 177-184, 1991
 A:Title: Molecular cloning and sequence analysis of bovine lactotransferrin.
 A:Reference number: S14110; MUID:91160550
 A:Accession: S14110
 A:Molecule type: mRNA
 A:Residues: 3-144, 'V', 146-339, 'A', 341-438, 'Y', 440-513, 'R', 515-708 <PI2>
 A:Cross-references: EMBL:X57084
 A:Accession: S18517
 A:Molecule type: protein
 A:Residues: 20-35; 82-114; 148-163; 'PP', 166-178, 'V', 'P', 183-190; 205-212; 230-239; 304-339
 R;Goodman, R.E.; Schanbacher, F.L.
 Biochem. Biophys. Res. Commun. 180, 75-84, 1991
 A:Title: Bovine lactoferrin mRNA: sequence, analysis, and expression in the mammary g
 A:Reference number: JTO595; MUID:92028986
 A:Accession: JTO595
 A:Molecule type: mRNA

```
A:Residues: 1-65,'PG',68-296,'S',298-339,'A',341-708 <GO>
A:Cross-references: GB:M63502
A>Note: The authors translated the codon CCG for residue 66 as Arg and TCT for residue 2
R:Mead, P.E.; Tweedie, J.W.
Nucleic Acids Res. 18, 7167, 1990
A:Title: cDNA and protein sequence of bovine lactoferrin.
A:Reference number: S13097; MUID:91088328
A:Accession: S13097
A:Molecule type: mRNA
A:Residues: 28-33,'DS',36-38,'P',40-708 <MEA>
A:Cross-references: EMBL:X54801
A:Accession: S18518
A:Molecule type: protein
A:Residues: 20-47:59-66;132-139;256-277;278,305-332;343-351;361-363;586,587-589;598-619
R:Mead, P.E.
Submitted to the EMBL Data Library, October 1990
A:Reference number: S13881
A:Accession: S13881
A:Molecule type: mRNA
A:Residues: 28-38,'P',40-86,'C',88-708 <ME3>
A:Cross-references: EMBL:X54801
R:Rejman, J.J.; Hegarty, H.M.; Hurley, W.L.
Comp. Biochem. Physiol. B 93, 929-934, 1989
A:Title: Purification and characterization of bovine lactoferrin from secretions of the
A:Reference number: PLO148; MUID:90031456
A:Accession: PLO148
A:Molecule type: protein
A:Residues: 20-27,'X',29-37,'X',39-54,'X',56-59 <REJ>
R:Beliamy, W.; Takase, M.; Yamauchi, K.; Wakabayashi, H.; Kawase, K.; Tomita, M.
Biochim. Biophys. Acta 1121, 130-136, 1992
A:Title: Identification of the bactericidal domain of lactoferrin.
A:Reference number: S21756; MUID:92287941
A:Accession: S21756
A:Molecule type: protein
A:Residues: 36-60 <BEL>
R:Shimazaki, K.; Tanaka, T.; Kon, H.; Oota, K.; Kawaguchi, A.; Maki, Y.; Sato, T.
J. Dairy Sci. 76, 946-955, 1993
A:Title: Separation and characterization of the C-terminal half molecule of bovine lacto
A:Reference number: A56659; MUID:93253156
A:Accession: A56659
A:Molecule type: protein
A:Residues: 20-25;302-308;359-366,'X',368-376,'X',378 <SHI>
A:Superfamily: transferrin; transferrin repeat homology
C:Keywords: duplication; glycoprotein; iron; iron binding; metalloprotein; milk
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-708/Product: lactotransferrin #status experimental <MAT>
F:20-355/Domain: transferrin repeat homology <TRH1>
F:36-60/Region: antimicrobial
F:359-696/Domain: transferrin repeat homology <TRH2>
F:28-64,134-217,176-192,179-200,189-202,250-264,367-399,377-390,424-703,444-666,476-551,
F:38-55/disulfide bonds: #status predicted
F:79,111,211,272/Binding site: iron (Asp, Tyr, Tyr, His) #status experimental
F:140/Binding site: carbonate (Arg) #status experimental
F:252,300,387,495,564/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:414,452,545,614/Binding site: iron (Arg, Tyr, Tyr, His) #status experimental
F:482/Binding site: carbonate (Arg) #status experimental

Query Match 75.0%; Score 45; DB 1; Length 708;
Best Local Similarity 75.0%; Pred. No. 2.4;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ARRARVVWAAVG 12
||| ||||| |||
Db 359 ARYTRVWCAVG 370

RESULT 4
JC2323
lactoferrin - goat
C:Species: Capra aegagrus hircus (domestic goat)
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 07-May-1999
C:Accession: JC2323

A:Residues: 1-65,'PG',68-296,'S',298-339,'A',341-708 <GO>
A:Cross-references: GB:M63502
A>Note: The authors translated the codon CCG for residue 66 as Arg and TCT for residue 2
R:Mead, P.E.; Tweedie, J.W.
Nucleic Acids Res. 18, 7167, 1990
A:Title: cDNA and protein sequence of bovine lactoferrin.
A:Reference number: S13097; MUID:91088328
A:Accession: S13097
A:Molecule type: mRNA
A:Residues: 28-33,'DS',36-38,'P',40-708 <MEA>
A:Cross-references: EMBL:X54801
A:Accession: S18518
A:Molecule type: protein
A:Residues: 20-47:59-66;132-139;256-277;278,305-332;343-351;361-363;586,587-589;598-619
R:Mead, P.E.
Submitted to the EMBL Data Library, October 1990
A:Reference number: S13881
A:Accession: S13881
A:Molecule type: mRNA
A:Residues: 28-38,'P',40-86,'C',88-708 <ME3>
A:Cross-references: EMBL:X54801
R:Rejman, J.J.; Hegarty, H.M.; Hurley, W.L.
Comp. Biochem. Physiol. B 93, 929-934, 1989
A:Title: Purification and characterization of bovine lactoferrin from secretions of the
A:Reference number: PLO148; MUID:90031456
A:Accession: PLO148
A:Molecule type: protein
A:Residues: 20-27,'X',29-37,'X',39-54,'X',56-59 <REJ>
R:Beliamy, W.; Takase, M.; Yamauchi, K.; Wakabayashi, H.; Kawase, K.; Tomita, M.
Biochim. Biophys. Acta 1121, 130-136, 1992
A:Title: Identification of the bactericidal domain of lactoferrin.
A:Reference number: S21756; MUID:92287941
A:Accession: S21756
A:Molecule type: protein
A:Residues: 36-60 <BEL>
R:Shimazaki, K.; Tanaka, T.; Kon, H.; Oota, K.; Kawaguchi, A.; Maki, Y.; Sato, T.
J. Dairy Sci. 76, 946-955, 1993
A:Title: Separation and characterization of the C-terminal half molecule of bovine lacto
A:Reference number: A56659; MUID:93253156
A:Accession: A56659
A:Molecule type: protein
A:Residues: 20-25;302-308;359-366,'X',368-376,'X',378 <SHI>
A:Superfamily: transferrin; transferrin repeat homology
C:Keywords: duplication; glycoprotein; iron; iron binding; metalloprotein; milk
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-708/Product: lactotransferrin #status experimental <MAT>
F:20-355/Domain: transferrin repeat homology <TRH1>
F:36-60/Region: antimicrobial
F:359-696/Domain: transferrin repeat homology <TRH2>
F:28-64,134-217,176-192,179-200,189-202,250-264,367-399,377-390,424-703,444-666,476-551,
F:38-55/disulfide bonds: #status predicted
F:79,111,211,272/Binding site: iron (Asp, Tyr, Tyr, His) #status experimental
F:140/Binding site: carbonate (Arg) #status experimental
F:252,300,387,495,564/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:414,452,545,614/Binding site: iron (Arg, Tyr, Tyr, His) #status experimental
F:482/Binding site: carbonate (Arg) #status experimental

Query Match 71.7%; Score 43; DB 2; Length 491;
Best Local Similarity 66.7%; Pred. No. 3.9;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ARRARVVWAAVG 12
||| ||||| |||
Db 58 ARQAQVANAAC 69

RESULT 6
A28438
lactoferrin precursor - mouse
N:Alternate names: lactotransferrin
C:Species: Mus musculus (house mouse)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A28438; A41205
R:Penetecost, B.T.; Teng, C.T.
J. Biol. Chem. 262, 10134-10139, 1987
A:Title: Lactotransferrin is the major estrogen inducible protein of mouse uterine se
A:Reference number: A92596; MUID:87280033
A:Accession: A28438
A:Molecule type: mRNA
A:Residues: 3-707 <PEN>
A:Cross-references: EMBL:J03298
R:Liu, Y.; Teng, C.T.
```

J. Biol. Chem. 266, 21880-21885, 1991
 A:Title: Characterization of estrogen-responsive mouse lactoferrin promoter.
 A:Reference number: A41205; MUID:92042099
 A:Accession: A41205
 A:Molecule type: DNA
 A:Residues: 1-15 <L10>
 A:Cross-references: GB:M74778
 C:Superfamily: transferrin; transferrin repeat homology
 C:Keywords: duplication; glycoprotein
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-707/Product: lactotransferrin #status predicted <MAT>
 F:358-695/Domain: transferrin repeat homology <TRH2>
 F:494/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 71.7%; Score 43; DB 1; Length 707;
 Best Local Similarity 66.7%; Pred. No. 5.5;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 ARRARVVAAGV 12
 I :||| I|||
 Db 358 ASKARVTWCAVG 369

RESULT 7
 JE0375
 carbonic anhydrase-related protein - human
 C:Species: Homo sapiens (man)
 C:Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 21-Jul-2000
 C:Accession: JE0375
 R:Beillingham, J.; Gregory-Evans, K.; Y.Gregory-Evans, C.
 Biochem. Biophys. Res. Commun. 253, 364-367, 1998
 A:Title: Sequence and tissue expression of a novel human carbonic anhydrase-related protein
 A:Reference number: JE0375; MUID:99097242
 A:Accession: JE0375
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-328 <BEL>
 A:Cross-references: GB:AF067662; NID:g3283385; PIDN:AC95689.1; PID:g3283386
 C:Superfamily: carbonic dehydratase; carbonic anhydrase homology
 F:45-303/Domain: carbonic anhydrase homology <CAH>

Query Match 66.7%; Score 40; DB 2; Length 328;
 Best Local Similarity 66.7%; Pred. No. 9.2;
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ARRARVVAAGV 12
 I ||| I|||I
 Db 9 APRALVLWAALG 20

RESULT 8
 H70526
 probable cytochrome P450 Rv0327c - Mycobacterium tuberculosis (strain H37RV)
 N:Contains: oxidoreductase (EC 1.-.-.-)
 C:Species: Mycobacterium tuberculosis
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
 C:Accession: H70526
 R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
 ; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
 Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A:Reference number: A70500; MUID:98295987
 A:Accession: H70526
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-449 <COL>
 A:Cross-references: GB:F96800; GB:AL123456; NID:g3261800; PIDN:CAB09576.1; PID:g2193948
 A:Experimental source: strain H37RV
 C:Genetics:

A:Gene: Rv0327c
 C:Superfamily: Bacillus cytochrome P450 CYP106; cytochrome P450 homology
 C:Keywords: chromoprotein; heme; iron; metalloprotein; oxidoreductase
 F:259-405/Domain: cytochrome P450 homology <P45>
 F:383/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 66.7%; Score 40; DB 1; Length 449;
 Best Local Similarity 70.0%; Pred. No. 12;
 Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 RARVVAAGV 12
 :|||:|||||
 Db 285 QSQVVAAGV 294

RESULT 9
 IDECRP
 replication initiation protein - Escherichia coli plasmids
 C:Species: Escherichia coli
 C:Date: 18-Dec-1981 #sequence_revision 17-Dec-1982 #text_change 16-Jul-1999
 C:Accession: A03602; B03602; S01773; S05591; A48662; C28378; I41109
 R:Rosen, J.; Ryder, T.; Inokuchi, H.; Ohtsubo, H.; Ohtsubo, E.
 Mol. Gen. Genet. 179, 527-537, 1980
 A:Title: Genes and sites involved in replication and incompatibility of an R100 plasmid
 A:Reference number: A93119; MUID:81074309
 A:Accession: A03602
 A:Molecule type: DNA
 A:Residues: 1-285 <NOL>
 A:Cross-references: GB:J01762; GB:J01761; GB:J01767; GB:J01768; NID:g151740; PIDN:AAA
 A:Experimental source: plasmid R100
 R:Rosen, J.; Ryder, T.; Ohtsubo, H.; Ohtsubo, E.
 Nature 290, 794-797, 1981
 A:Title: Role of RNA transcripts in replication incompatibility and copy number control
 A:Reference number: A93253; MUID:81173118
 A:Accession: B03602
 A:Molecule type: DNA
 A:Residues: 1-77 <ROS>
 A:Experimental source: plasmid R1
 R:Dong, X.; Womble, D.D.; Rownd, R.H.
 J. Mol. Biol. 202, 495-509, 1988
 A:Title: In-vivo studies on the cis-acting replication initiator protein of IncFII pl
 A:Reference number: S01773; MUID:89011975
 A:Accession: S01773
 A:Molecule type: DNA
 A:Residues: 1-285 <DON>
 A:Cross-references: EMBL:X12776; NID:g42707; PIDN:CAA31263.1; PID:g581213
 A:Experimental source: plasmid NRI
 R:Masai, H.; Arai, K.I.
 Nucleic Acids Res. 16, 6493-6514, 1988
 A:Title: RepA protein- and oriR-dependent initiation of R1 plasmid replication: ident
 A:Reference number: S05591; MUID:88289416
 A:Accession: S05591
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 270-285 <MAS>
 A:Cross-references: EMBL:X12587; NID:g41107; PIDN:CAA31100.1; PID:g41108
 A:Experimental source: plasmid R1
 R:Jiang, T.; Min, Y.N.; Liu, W.; Womble, D.D.; Rownd, R.H.
 J. Bacteriol. 175, 5350-5358, 1993
 A:Title: Insertion and deletion mutations in the repA region of the IncFII plasmid N
 A:Reference number: A48662; MUID:93374828
 A:Accession: A48662
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 252-285 <JIA>
 A:Experimental source: plasmid NRI
 R:Dong, X.; Womble, D.D.; Rownd, R.H.
 J. Bacteriol. 169, 5353-5363, 1987
 A:Title: Transcriptional pausing in a region important for plasmid NRI replication co
 A:Reference number: A28378; MUID:88058738
 A:Accession: C28378
 A:Status: preliminary

A:Molecule type: DNA
A:Residues: 1-50 <D02>
R:Womble, D.D.; Sampathkumar, P.; Easton, A.M.; Luckow, V.A.; Rownd, R.H.
J. Mol. Biol. 181, 395-410, 1985
A:Title: Transcription of the replication control region of the IncFII R-plasmid NR1 in
A:Reference number: I41106; MUID:85160860
A:Accession: I41109
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-285 <RES>
A:Cross-references: EMBL:X02302; NID:g42132; PIDN:CAA26168.1; PID:g581144
C:Genetics:
A:Gene: repA1
A:Genome: plasmid
A:Start codon: GTG
C:Superfamily: repl protein
C:Keywords: plasmid copy control

Query Match 65.0%; Score 39; DB 1; Length 285;
Best Local Similarity 87.5%; Pred. No. 12;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARARVVW 8
| | | | |
Db 168 ARSRVW 175

RESULT 10
I64780
replication-associated protein A1 - Escherichia coli
C:Species: Escherichia coli
C:Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 26-Aug-1999
A:Accession: I64780; I41110
R:Ohtsubo, H.; Ryder, T.B.; Maeda, Y.; Armstrong, K.A.; Ohtsubo, E.
Adv. Biophys. 21, 115-133, 1986
A:Title: DNA replication of the resistance plasmid R100 and its control.
A:Reference number: I51821; MUID:86319522
A:Accession: I64780
A:Status: translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-285 <RES>
A:Cross-references: GB:M26840; NID:gl51770; PIDN:AAA26067.1; PID:gl51773
R:Womble, D.D.; Sampathkumar, P.; Easton, A.M.; Luckow, V.A.; Rownd, R.H.
J. Mol. Biol. 181, 395-410, 1985
A:Title: Transcription of the replication control region of the IncFII R-plasmid NR1 in
A:Reference number: I41106; MUID:85160860
A:Accession: I41110
A:Status: translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 'M',166-285 <RE2>
A:Cross-references: EMBL:X02302; NID:g42132; PIDN:CAA26169.1; PID:g581145
C:Superfamily: repl protein

Query Match 65.0%; Score 39; DB 2; Length 285;
Best Local Similarity 87.5%; Pred. No. 12;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARARVVW 8
| | | | |
Db 168 ARSRVW 175

RESULT 11
F72563
hypothetical protein APE1792 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
A:Accession: F72563
R:Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999

A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aero
A:Reference number: A72450; MUID:99310339
A:Accession: F72563
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-176 <KAW>
A:Cross-references: DDBJ:AP000062; NID:g5105244; PIDN:BAA80795.1; PID:g5105482
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE1792
C:Superfamily: conserved hypothetical protein MJ1568

Query Match 63.3%; Score 38; DB 2; Length 176;
Best Local Similarity 70.0%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 RARVVWAAVG 12
| | | | |
Db 72 RPRVWVG 81

RESULT 12
T02260
beta-fructofuranosidase (EC 3.2.1.26) - maize (fragment)
N:Alternate names: invertase
C:Species: Zea mays (maize)
C:Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 22-Jun-1999
C:Accession: T02260
R:Xu, J.; Almira, E.C.; Avigne, W.T.; McCarty, D.R.; Koch, K.E.
submitted to the EMBL Data Library, July 1995
A:Description: Molecular characterization and differential expression of an invertase
A:Reference number: Z14645
A:Accession: T02260
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-509 <XUJ>
A:Cross-references: EMBL:U31451; NID:g951165; PIDN:AAA74584.1; PID:g951166
A:Experimental source: strain merit; root tip
C:Function:
A:Description: catalyzes hydrolyzation of terminal non-reducing beta-D-fructofuranosi
C:Superfamily: beta-fructofuranosidase
C:Keywords: glycosidase; hydrolase

Query Match 63.3%; Score 38; DB 2; Length 509;
Best Local Similarity 58.3%; Pred. No. 32;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ARARVVWAAVG 12
| | | | |
Db 248 AKRRRVWVG 259

RESULT 13
F72507
hypothetical protein APE2036 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jun-2000
C:Accession: F72507
R:Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Ta
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aero
A:Reference number: A72450; MUID:99310339
A:Accession: F72507
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-108 <KAW>
A:Cross-references: DDBJ:AP000063; NID:g5105654; PIDN:BAA81046.1; PID:d1044832; PID:g
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE2036

C:Superfamily: Aeropyrum pernix hypothetical protein APE2036

Query Match 61.7%; Score 37; DB 2; Length 108;
Best Local Similarity 66.7%; Pred. No. 11;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ARARVWVA 9
| | | | |
DB 63 AERARIWVS 71

RESULT 14

T12816
hypothetical protein yonB - Bacillus subtilis phage SPBc2
C:Species: Bacillus subtilis phage SPBc2
C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 20-Jun-2000
C:Accession: T12816; D69913
R:Lazarevic, V.; Duesterhoeft, A.; Solido, B.; Hilbert, H.; Mauer, C.; Karamata, D.
submitted to the EMBL Data Library, August 1997
A:Description: The complete nucleotide sequence of the Bacillus subtilis SPBc2 prophage
A:Reference number: Z17583

A:Accession: T12816

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-338 <LAZ>

A:Cross-references: EMBL:AF020713; NID:g3025478; PID:g3025530; PIDN:AC13025.1
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berth
C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F.
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetel
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; MUID:98044033

A:Accession: D69913

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-338 <RUN>

A:Cross-references: GB:299115; GB:AL009126; NID:g2634478; PIDN:CAB14033.1; PID:g2634535

A:Experimental source: strain 168

C:Genetics:

A:Gene: yonB

C:Superfamily: Bacillus subtilis phage SPBc2 hypothetical protein yonB

Query Match 61.7%; Score 37; DB 2; Length 338;
Best Local Similarity 50.0%; Pred. No. 33;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 3 RARVWVAAG 12
| | | | |
DB 95 KAKVWSATG 104

RESULT 15

A75503
hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C:Accession: A75503
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathavan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896

A:Accession: A75503

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-397 <WHI>

A:Cross-references: GB:AE001915; GB:AE000513; NID:g6458262; PIDN:AAF10152.1; PID:g645

A:Experimental source: strain R1

C:Genetics:

A:Gene: DR0571

A:Map position: 1

Query Match 61.7%; Score 37; DB 2; Length 397;

Best Local Similarity 72.7%; Pred. No. 38;

Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 RRARVWVAAG 12

| | | | |

DB 19 RRAGSVWVAHV 29

Search completed: December 31, 2001, 08:52:49

Job time: 43 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 31, 2001, 08:52:06 ; Search time 10.12 Seconds
(without alignments)
43.476 Million cell updates/sec

Title: US-09-508-095-22
Perfect score: 60
Sequence: 1 ARRARVVAAGV 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues
Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	56	93.3	708	1 TRFL_BUBBU	O77698 bubalus bub
2	56	93.3	711	1 TRFL_HUMAN	P02788 homo sapien
3	51	85.0	695	1 TRFL_HORSE	O77811 equus caball
4	49	81.7	704	1 TRFL_PIG	P14632 sus scrofa
5	48	80.0	708	1 TRFL_CAMD	Q9cum0 camelus dro
6	45	75.0	708	1 TRFL_BOVIN	P24627 bos taurus
7	44	73.3	708	1 TRFL_CAPIH	Q29477 capra hircu
8	43	71.7	707	1 TRFL_MOUSE	P08071 mus musculu
9	40	66.7	328	1 CAHB_HUMAN	O75493 homo sapien
10	40	66.7	449	1 C13A_MYCTU	O08447 mycobacteri
11	39	65.0	285	1 REP2_ECOLI	P03066 escherichia
12	38	63.3	176	1 YH92_AERPE	Q9yb03 aeropyrum p
13	37	61.7	368	1 GALT_HUMAN	O60755 homo sapien
14	37	61.7	651	1 BGLR_HUMAN	P08236 homo sapien
15	36	60.0	328	1 CAHB_SHEEP	Q95203 ovis aries
16	36	60.0	501	1 VLI_PAPVE	P11326 european el
17	36	60.0	507	1 VLI_HPV15	Q05137 human papil
18	36	60.0	513	1 VLI_PAPVD	P03104 deer papill
19	36	60.0	771	1 R1R1_VACCC	P20503 vaccinia vi
20	36	60.0	771	1 R1R1_VACCV	P12848 vaccinia vi
21	36	60.0	771	1 R1R1_VARV	P32984 variola vir
22	36	60.0	895	1 GNDS_RAT	Q03385 rattus norv
23	36	60.0	914	1 GNDS_HUMAN	Q22967 homo sapien
24	35	58.3	495	1 VLI_BFV1	P03103 bovine pap
25	35	58.3	497	1 VLI_BFV2	P06458 bovine pap
26	35	58.3	499	1 VLI_HPV33	P06416 human papil
27	35	58.3	499	1 VLI_HPV53	Q05113 human papil
28	35	58.3	502	1 VLI_HPV42	P27233 human papil
29	35	58.3	503	1 VLI_HPV26	P36735 human papil
30	35	58.3	503	1 VLI_HPV32	P36737 human papil
31	35	58.3	503	1 VLI_HPV66	Q80961 human papil
32	35	58.3	504	1 VLI_HPV31	P17388 human papil
33	35	58.3	505	1 VLI_HPV07	Q05136 human papil

34	35	58.3	505	1 VLI_HPV39	P24838 human papil
35	35	58.3	505	1 VLI_HPV40	P36740 human papil
36	35	58.3	505	1 VLI_HPV68	P54669 human papil
37	35	58.3	505	1 VLI_HPVME	P27964 human papil
38	35	58.3	508	1 VLI_HPV30	Q02515 human papil
39	35	58.3	524	1 VLI_HPV58	P26335 human papil
40	35	58.3	531	1 VLI_HPV16	P03101 human papil
41	35	58.3	534	1 VLI_HPV56	P36741 human papil
42	35	58.3	539	1 VLI_HPV45	P36741 human papil
43	35	58.3	560	1 JI60_HORVU	Q00531 hordeum vul
44	35	58.3	704	1 ICA_PIG	Q29545 sus scrofa
45	34	56.7	366	1 YGDE_ECOLI	P32066 escherichia

ALIGNMENTS

```

RESULT 1
TRFL_BUBBU
ID TRFL_BUBBU STANDARD; PRT; 708 AA.
AC O77698;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE LACTOTRANSFERRIN PRECURSOR (LACTOFERRIN).
GN LTF.
OS Bubalus bubalis (Domestic water buffalo).
OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bubalus.
OX NCBI_TaxID=89462;
RN [1]
RP SEQUENCE FROM N.A.
RA Paramasivam M., Thattaliyath B.D., Kumar A., Srinivasan A.,
RA Singh T.P.;
RT "cDNA sequence of Buffalo lactoferrin.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RX MEDLINE=20003130; PubMed=10531476;
RA Karthikeyan S., Paramasivam M., Yadav S., Srinivasan A., Singh T.P.;
RT "Structure of buffalo lactoferrin at 2.5-A resolution using crystals
RT grown at 303 K shows different orientations of the N and C lobes.";
RL Acta Crystallogr. D 55:1805-1813(1999).
CC -!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH
CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING
CC OF AN ANION, USUALLY BICARBONATE.
CC -!- SUBUNIT: MONOMER.
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.
CC -!- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ005203; CAA06441.1;
DR PDB; ICE2; 19-MAR-99.
DR InterPro; IPR001156; Transferrin.
DR Pfam; PF00405; transferrin; 2.
DR PRINTS; PR00422; TRANSFERRIN.
DR SMART; SM00094; TR_FER; 2.
DR PROSITE; PS00205; TRANSFERRIN_1; 2.
DR PROSITE; PS00206; TRANSFERRIN_2; 2.
DR PROSITE; PS00207; TRANSFERRIN_3; 2.
DR Iron transport; Glycoprotein; Metal-binding; Repeat; Signal;
KW 3D-structure.
FT SIGNAL 1 19
FT CHAIN 20 708 LACTOTRANSFERRIN.

```

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

CC EMBL; AJ010930; CAA09407.1; -
 CC PDB; 1B1X; 02-DEC-98.
 CC PDB; 1B7U; 02-FEB-99.
 CC PDB; 1B7Z; 02-FEB-99.
 CC InterPro; IPR001156; Transferrin.
 CC Pfam; PF00405; transferrin; 2.
 CC PRINTS; PR00422; TRANSFERRIN.
 CC SMART; SM00094; TR_FER; 2.
 CC PROSITE; PS00205; TRANSFERRIN_1; 2.
 CC PROSITE; PS00206; TRANSFERRIN_2; 2.
 CC PROSITE; PS00207; TRANSFERRIN_3; 1.
 CC Iron transport; Glycoprotein; Metal-binding; Repeat; Signal;
 CC 3D-structure.
 CC NON_TER 1 1
 CC SIGNAL <1 6
 CC CHAIN 7 695 LACTOTRANSFERRIN.
 CC REPEAT 7 350 1.
 CC REPEAT 351 695 2.
 CC DISULFID 15 51
 CC DISULFID 25 42
 CC DISULFID 121 204
 CC DISULFID 163 179
 CC DISULFID 166 189
 CC DISULFID 176 187
 CC DISULFID 237 251
 CC DISULFID 354 386
 CC DISULFID 364 377
 CC DISULFID 411 690
 CC DISULFID 431 653
 CC DISULFID 463 538
 CC DISULFID 487 681
 CC DISULFID 497 511
 CC DISULFID 508 521
 CC DISULFID 579 593
 CC DISULFID 631 636
 CC METAL 66 66 IRON 1 (BY SIMILARITY).
 CC METAL 98 98 IRON 1 (BY SIMILARITY).
 CC METAL 198 198 IRON 1 (BY SIMILARITY).
 CC METAL 259 259 IRON 1 (BY SIMILARITY).
 CC METAL 401 401 IRON 2 (BY SIMILARITY).
 CC METAL 439 439 IRON 2 (BY SIMILARITY).
 CC METAL 532 532 IRON 2 (BY SIMILARITY).
 CC METAL 601 601 IRON 2 (BY SIMILARITY).
 CC BINDING 127 127 ANION (BY SIMILARITY).
 CC BINDING 469 469 ANION (BY SIMILARITY).
 CC CARBOHYD 143 143 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 287 287 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 482 482 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC SEQUENCE 695 AA; 75991 MW; 07BB84D50E1B165D CRC64;
 GN LTF.
 OS Sus scrofa (Pig).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 CC NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-92379101; PubMed-1511016;
 RA Lyndon J.P., O'Malley B.R., Saucedo O., Lee T., Headon D.R.,
 RA Conneely O.M.;
 RT "Nucleotide and primary amino acid sequence of porcine lactoferrin.";
 RL Biochim. Biophys. Acta 1132:97-99(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-92367939; PubMed-1503259;
 RA Alexander L.J., Levine W.B., Teng C.T., Beattie C.W.;
 RT "Cloning and sequencing of the porcine lactoferrin cDNA.";
 RL Anlm. Genet. 23:251-256(1992).
 RN [3]
 RP SEQUENCE OF 20-49.
 RX MEDLINE-90105538; PubMed-2605266;
 RA Hutchens T.W., Magnuson J.S., Yip T.-T.;
 RT "Rapid purification of porcine colostrum whey lactoferrin by affinity
 chromatography on single-stranded DNA-agarose. Characterization,
 amino acid composition and N-terminal amino acid sequence.";
 RL Biochim. Biophys. Acta 999:323-329(1989).
 CC -!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH
 CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING
 CC OF AN ANION, USUALLY BICARBONATE.
 CC -!- SUBUNIT: MONOMER.
 CC -!- SUBCELLULAR LOCATION: SECRETED.
 CC -!- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.
 CC -!- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

Query Match 85.0%; Score 51; DB 1; Length 695;
 Best Local Similarity 83.3%; Pred. No. 0.086;
 Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ARRRVVAWG 12
 III IIII III
 Db 346 ARRRVVAWG 357

RESULT 4
 TRFL_PIG STANDARD; PRT; 704 AA.
 AC P14632; Q29557;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE LACTOTRANSFERRIN PRECURSOR (LACTOFERRIN).


```
FT METAL 107 107 IRON 1 (BY SIMILARITY).
FT METAL 207 207 IRON 1 (BY SIMILARITY).
FT METAL 268 268 IRON 1 (BY SIMILARITY).
FT METAL 410 410 IRON 2 (BY SIMILARITY).
FT METAL 448 448 IRON 2 (BY SIMILARITY).
FT METAL 541 541 IRON 2 (BY SIMILARITY).
FT METAL 610 610 ANION (BY SIMILARITY).
FT BINDING 136 136 ANION (BY SIMILARITY).
FT CARBOHYD 478 478 ANION (BY SIMILARITY).
FT CARBOHYD 385 385 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 491 491 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 12 12 G -> W (IN REF. 1).
FT CONFLICT 46 46 RRT -> TTR (IN REF. 3).
FT CONFLICT 51 51 M -> I (IN REF. 1).
FT CONFLICT 85 85 D -> G (IN REF. 1).
FT CONFLICT 121 121 MISSING (IN REF. 2).
FT CONFLICT 132 132 T -> I (IN REF. 1).
FT CONFLICT 284 284 E -> S (IN REF. 1).
FT CONFLICT 573 573 E -> Q (IN REF. 1).
FT CONFLICT 590 590 D -> N (IN REF. 1).
FT CONFLICT 625 625 V -> M (IN REF. 1).
FT CONFLICT 662 662 V -> C (IN REF. 1).
FT CONFLICT 686 704 NLKQCSVSPLLACAFMMR -> T (IN REF. 1).
SQ SEQUENCE 704 AA; 77625 MW; 93261EFD608AD358 CRC64;

Query Match 81.7%; Score 49; DB 1; Length 704;
Best Local Similarity 75.0%; Pred. No. 0.2;
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ARRAVVWAAVG 12
   1111111111
DB 355 ARQAVVWCAVG 366

RESULT 5
TRFL_CAMDR STANDARD; PRT; 708 AA.
AC Q9TUM0; Q9MZS5;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DE LACTOTRANSFERRIN PRECURSOR (LACTOFERRIN).
GN LTF.
OS Camelus dromedarius (Dromedary) (Arabian camel).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Camelus.
OX NCBI_TaxID=9838;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SONALI; TISSUE=Lactating mammary gland;
RA Kappeler S.R.; Ackermann M.; Farah Z.; Puhon Z.;
RT "Sequence analysis of camel (Camelus dromedarius) lactoferrin.";
RL Int. Dairy J. 9:481-486(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RA Paramasivam M.; Srinivasan A.; Singh R.; Sahani M.S.; Singh T.P.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH
CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING
CC OF AN ANION, USUALLY BICARBONATE (BY SIMILARITY).
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.
CC -1- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
```

```
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ131674; CAB53387.1; -.
DR EMBL; AF165879; AAF82241.1; -.
DR HSSP; O77811; 1BIX.
DR InterPro; IPR001156; Transferrin.
DR Pfam; PF00405; transferrin; 2.
DR PRINTS; PR00422; TRANSFERRIN.
DR SMART; SM00094; TR_FER; 2.
DR PROSITE; PS00205; TRANSFERRIN_1; 3.
DR PROSITE; PS00206; TRANSFERRIN_2; 2.
DR PROSITE; PS00207; TRANSFERRIN_3; 2.
KW Iron transport; Glycoprotein; Metal-binding; Repeat; Signal.
FT SIGNAL 1 19 BY SIMILARITY.
FT CHAIN 20 708 LACTOTRANSFERRIN.
FT REPEAT 20 363 1.
FT REPEAT 364 708 2.
FT DISULFID 28 64 BY SIMILARITY.
FT DISULFID 38 55 BY SIMILARITY.
FT DISULFID 134 217 BY SIMILARITY.
FT DISULFID 176 192 BY SIMILARITY.
FT DISULFID 189 200 BY SIMILARITY.
FT DISULFID 250 264 BY SIMILARITY.
FT DISULFID 357 399 BY SIMILARITY.
FT DISULFID 377 390 BY SIMILARITY.
FT DISULFID 424 703 BY SIMILARITY.
FT DISULFID 444 666 BY SIMILARITY.
FT DISULFID 476 551 BY SIMILARITY.
FT DISULFID 500 694 BY SIMILARITY.
FT DISULFID 510 524 BY SIMILARITY.
FT DISULFID 521 534 BY SIMILARITY.
FT DISULFID 592 606 BY SIMILARITY.
FT DISULFID 644 649 BY SIMILARITY.
FT METAL 79 79 IRON 1 (BY SIMILARITY).
FT METAL 111 111 IRON 1 (BY SIMILARITY).
FT METAL 211 211 IRON 1 (BY SIMILARITY).
FT METAL 272 272 IRON 1 (BY SIMILARITY).
FT METAL 414 414 IRON 2 (BY SIMILARITY).
FT METAL 452 452 IRON 2 (BY SIMILARITY).
FT METAL 545 545 IRON 2 (BY SIMILARITY).
FT METAL 614 614 IRON 2 (BY SIMILARITY).
FT BINDING 140 140 ANION (BY SIMILARITY).
FT BINDING 482 482 ANION (BY SIMILARITY).
FT CARBOHYD 252 252 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 385 385 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 537 537 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 594 594 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 261 261 F -> S (IN REF. 2).
FT CONFLICT 304 304 G -> A (IN REF. 2).
FT CONFLICT 330 330 S -> P (IN REF. 2).
FT CONFLICT 492 494 LIS -> PLF (IN REF. 2).
FT CONFLICT 506 506 L -> F (IN REF. 2).
FT CONFLICT 609 609 A -> P (IN REF. 2).
FT CONFLICT 642 642 R -> Q (IN REF. 2).
SQ SEQUENCE 708 AA; 77211 MW; 0B0C175A0B69D430 CRC64;

Query Match 80.0%; Score 48; DB 1; Length 708;
Best Local Similarity 81.8%; Pred. No. 0.3;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 RRAVVWAAVG 12
   1111111111
DB 360 RRAQVWCAVG 370

RESULT 6
TRFL_BOVIN STANDARD; PRT; 708 AA.
ID TRFL_BOVIN
AC P24627; Q29629; Q9M2Y3;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
```

DE LACTOTRANSFERRIN PRECURSOR (LACTOFERRIN) [CONTAINS: LACTOFERRICIN B
DE (LFCIN B)].
GN LTF.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RN SEQUENCE FROM N.A.
RN TISSUE-Submaxillary gland;
RX MEDLINE-91160550; PubMed=2001696;
RA Pierce A., Colavizza D., Benaissa M., Maes P., Tartar A.,
RA Montreuil J., Spik G.;
RT "Molecular cloning and sequence analysis of bovine lactotransferrin";
RL Eur. J. Biochem. 196;177-184(1991).
RN [2]
RN SEQUENCE FROM N.A.
RX MEDLINE-92028986; PubMed=1718281;
RA Goodman R.E., Schanbacher F.L.;
RT "Bovine lactoferrin mRNA: sequence, analysis, and expression in the
RT mammary gland";
RL Biochem. Biophys. Res. Commun. 180:75-84(1991).
RN [3]
RN SEQUENCE FROM N.A.
RN TISSUE-Lung;
RX Tsang T.C., Burns D.K., Wang F., Pan Y.C.E., Schmidt A.M., Stern D.;
RT "Cloning of a 80-KD advanced glycosylation end product (AGE) binding
RT protein from bovine lung";
RL FASEB J. 6:233-233(1991).
RN [4]
RN SEQUENCE FROM N.A.
RN TISSUE-Blood, and Mammary gland;
RX MEDLINE-94266164; PubMed=8206385;
RA Seyfert H.M., Tuckoricz A., Interthal H., Koczan D., Hobom G.;
RT "Structure of the bovine lactoferrin-encoding gene and its promoter";
RL Gene 143:265-269(1994).
RN [5]
RN SEQUENCE FROM N.A.
RA Nakamura I., Shimazaki K., Yagi Y., Watanabe A.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN [6]
RN SEQUENCE OF 20-59.
RX MEDLINE-90031466; PubMed=2805645;
RA Rejman J.J., Hegarty H.M., Hurley W.L.;
RT "Purification and characterization of bovine lactoferrin from
RT secretions of the involuting mammary gland: identification of
RT multiple molecular weight forms";
RL Comp. Biochem. Physiol. 93B:929-934(1989).
RN [7]
RN X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).
RX MEDLINE-98062367; PubMed=9398529;
RA Moore S.A., Anderson B.F., Groom C.R., Haridas M., Baker E.N.;
RT "Three-dimensional structure of diferric bovine lactoferrin at 2.8-A
RT resolution";
RL J. Mol. Biol. 274:222-236(1997).
RN [8]
RN STRUCTURE BY NMR OF 36-60.
RX MEDLINE-98190007; PubMed=9521752;
RA Wang P.M., Zhou N., Shan X., Arrowsmith C.H., Vogel H.J.;
RT "Three-dimensional solution structure of lactoferricin B, an
RT antimicrobial peptide derived from bovine lactoferrin";
RL Biochemistry 37:4288-4298(1998).
CC -1- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH
CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING
CC OF AN ANION, USUALLY BICARBONATE.
CC -1- FUNCTION: LACTOFERRICIN B IS AN ANTIMICROBIAL PEPTIDE.
CC -1- SUBUNIT: MONOMER.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.
CC -1- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X57084; CAA04366.1; -
DR EMBL; M63502; AAA30617.1; -
DR EMBL; L08604; AAA30609.1; -
DR EMBL; L19993; AAA21722.1; -
DR EMBL; L19982; AAA21722.1; JOINED.
DR EMBL; L19983; AAA21722.1; JOINED.
DR EMBL; L19984; AAA21722.1; JOINED.
DR EMBL; L19985; AAA21722.1; JOINED.
DR EMBL; L19986; AAA21722.1; JOINED.
DR EMBL; L19988; AAA21722.1; JOINED.
DR EMBL; L19989; AAA21722.1; JOINED.
DR EMBL; L19990; AAA21722.1; JOINED.
DR EMBL; L19991; AAA21722.1; JOINED.
DR EMBL; L19992; AAA21722.1; JOINED.
DR EMBL; AB046664; BAB03470.1; -
DR PIR; SI4674; TFBOL.
DR PIR; JT0595; JT0595.
DR PDB; 1BLF; 03-DEC-97.
DR PDB; 1LFC; 18-NOV-98.
DR GlycoSuiteDB; P24627; -
DR InterPro; IPR001156; Transferrin.
DR Pfam; PF00405; transferrin; 2.
DR PRINTS; PR00422; TRANSFERRIN.
DR SMART; SM00094; TR_FER; 2.
DR PROSITE; PS00205; TRANSFERRIN_1; 2.
DR PROSITE; PS00206; TRANSFERRIN_2; 2.
DR PROSITE; PS00207; TRANSFERRIN_3; 2.
KW Iron transport; Glycoprotein; Metal-binding; Repeat; Signal;
KW Antibiogenic; 3D-structure.
FT SIGNAL 1 19
FT CHAIN 20 708 LACTOTRANSFERRIN.
FT PEPTIDE 36 60 LACTOFERRICIN B.
FT REPEAT 20 363 1.
FT REPEAT 364 708 2.
FT DISULFID 28 64 BY SIMILARITY.
FT DISULFID 38 55 BY SIMILARITY.
FT DISULFID 134 217 BY SIMILARITY.
FT DISULFID 176 192 BY SIMILARITY.
FT DISULFID 189 200 BY SIMILARITY.
FT DISULFID 250 264 BY SIMILARITY.
FT DISULFID 367 399 BY SIMILARITY.
FT DISULFID 377 390 BY SIMILARITY.
FT DISULFID 424 703 BY SIMILARITY.
FT DISULFID 444 666 BY SIMILARITY.
FT DISULFID 476 551 BY SIMILARITY.
FT DISULFID 500 694 BY SIMILARITY.
FT DISULFID 510 524 BY SIMILARITY.
FT DISULFID 521 534 BY SIMILARITY.
FT DISULFID 592 606 BY SIMILARITY.
FT DISULFID 644 649 BY SIMILARITY.
FT METAL 79 79 IRON 1.
FT METAL 111 111 IRON 1.
FT METAL 211 211 IRON 1.
FT METAL 272 272 IRON 1.
FT METAL 414 414 IRON 2.
FT METAL 452 452 IRON 2.
FT METAL 545 545 IRON 2.
FT METAL 614 614 IRON 2.
FT BINDING 140 140 ANION (POTENTIAL).
FT BINDING 482 482 ANION (POTENTIAL).
FT CARBOHYD 252 252 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 387 387 N-LINKED (GLCNAC. . .).
FT CARBOHYD 495 495 N-LINKED (GLCNAC. . .).
FT CARBOHYD 564 564 N-LINKED (GLCNAC. . .).
FT CONFLICT 63 63 E -> A (IN REF. 4).
FT CONFLICT 66 67 RA -> PG (IN REF. 2).

```
FT CONFLICT 145 145 I -> V (IN REF. 1 AND 4).
FT CONFLICT 164 165 LQ -> PP (IN REF. 1).
FT CONFLICT 264 264 C -> Y (IN REF. 4).
FT CONFLICT 273 273 A -> P (IN REF. 4).
FT CONFLICT 281 281 G -> A (IN REF. 4).
FT CONFLICT 291 291 S -> R (IN REF. 4).
FT CONFLICT 297 297 F -> S (IN REF. 2).
FT CONFLICT 340 340 G -> A (IN REF. 1).
FT CONFLICT 418 418 I -> V (IN REF. 4).
FT CONFLICT 439 439 H -> Y (IN REF. 1).
FT CONFLICT 459 459 K -> R (IN REF. 4).
FT CONFLICT 514 514 A -> R (IN REF. 1).
FT CONFLICT 632 632 H -> R (IN REF. 5).
SQ SEQUENCE 708 AA; 78056 MW; C6FD/FC15D68E93F CRC64;

Query Match 75.0%; Score 45; DB 1; Length 708;
Best Local Similarity 75.0%; Pred. No. 1.1;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ARRVRVWAAVG 12
Db 359 ARYTRVWCAVG 370

RESULT 7
TRFL_CAPHI STANDARD; PRT; 708 AA.
AC Q29477; Q29479;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE LACTOTRANSFERRIN PRECURSOR (LACTOFERRIN).
GN LTF.
OS Capra hircus (Goat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Capra.
OX NCBI_TaxID=9925;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RX MEDLINE=94380047; PubMed=8093048;
RA le Provost F., Nocart M., Guerin G., Martin P.;
RT "Characterization of the goat lactoferrin cDNA. Assignment of the
RL Biochem. Biophys. Res. Commun. 203:1324-1332(1994).
CC -!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH
CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING
CC OF AN ANION, USUALLY BICARBONATE.
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.
CC -!- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U53857; AAA97958.1; -
DR EMBL; X78902; CAA55517.1; -
DR InterPro; IPR001156; Transferrin.
DR Pfam; PF00405; transferrin; 2.
DR PRINTS; PR00422; TRANSFERRIN.
```

```
DR SMART; SM00094; TR_FER; 2.
DR PROSITE; PS00205; TRANSFERRIN_1; 2.
DR PROSITE; PS00206; TRANSFERRIN_2; 2.
DR PROSITE; PS00207; TRANSFERRIN_3; 2.
KW Iron transport; Glycoprotein; Metal-binding; Repeat; Signal.
FT SIGNAL 1 19 BY SIMILARITY.
FT CHAIN 20 708 LACTOTRANSFERRIN.
FT REPEAT 20 363 1.
FT REPEAT 364 708 2.
FT DISULFID 28 64 BY SIMILARITY.
FT DISULFID 38 55 BY SIMILARITY.
FT DISULFID 134 217 BY SIMILARITY.
FT DISULFID 176 192 BY SIMILARITY.
FT DISULFID 189 200 BY SIMILARITY.
FT DISULFID 250 264 BY SIMILARITY.
FT DISULFID 367 399 BY SIMILARITY.
FT DISULFID 377 390 BY SIMILARITY.
FT DISULFID 424 703 BY SIMILARITY.
FT DISULFID 444 666 BY SIMILARITY.
FT DISULFID 476 551 BY SIMILARITY.
FT DISULFID 500 694 BY SIMILARITY.
FT DISULFID 510 524 BY SIMILARITY.
FT DISULFID 521 534 BY SIMILARITY.
FT DISULFID 592 606 BY SIMILARITY.
FT DISULFID 644 649 BY SIMILARITY.
FT METAL 79 79 IRON 1 (BY SIMILARITY).
FT METAL 111 111 IRON 1 (BY SIMILARITY).
FT METAL 211 211 IRON 1 (BY SIMILARITY).
FT METAL 272 272 IRON 1 (BY SIMILARITY).
FT METAL 414 414 IRON 2 (BY SIMILARITY).
FT METAL 452 452 IRON 2 (BY SIMILARITY).
FT METAL 545 545 IRON 2 (BY SIMILARITY).
FT METAL 614 614 IRON 2 (BY SIMILARITY).
FT BINDING 140 140 ANION (BY SIMILARITY).
FT BINDING 482 482 ANION (BY SIMILARITY).
FT CARBOHYD 252 252 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 300 300 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 387 387 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 495 495 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 564 564 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 56 56 I -> V (IN REF. 2).
FT CONFLICT 88 88 L -> R (IN REF. 2).
FT CONFLICT 124 124 Q -> K (IN REF. 2).
FT CONFLICT 154 154 F -> P (IN REF. 2).
FT CONFLICT 304 304 S -> R (IN REF. 2).
FT CONFLICT 414 414 D -> G (IN REF. 2).
SQ SEQUENCE 708 AA; 77358 MW; F2EDA3C83539960D CRC64;

Query Match 73.3%; Score 44; DB 1; Length 708;
Best Local Similarity 75.0%; Pred. No. 1.6;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ARRVRVWAAVG 12
Db 359 ARYTRVWCAVG 370

RESULT 8
TRFL_MOUSE STANDARD; PRT; 707 AA.
AC P08071; P70690; Q61799;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE LACTOTRANSFERRIN PRECURSOR (LACTOFERRIN).
GN LTF.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
```

RC TISSUE=Uterus;
RA MEDLINE=87280033; PubMed=3611056;
RX Pentecost B.T., Teng C.T.;
RT "Lactoferrin is the major estrogen inducible protein of mouse
RL uterine secretions";
RN J. Biol. Chem. 262:10134-10139(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Uterus;
RA Morishiki K.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RX MEDLINE=92042099; PubMed=1939212;
RA Liu Y., Teng C.T.;
RT "Characterization of estrogen-responsive mouse lactoferrin promoter";
RL J. Biol. Chem. 266:21880-21885(1991).
CC -1- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH
CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING
CC OF AN ANION, USUALLY BICARBONATE.
CC -1- SUBUNIT: MONOMER.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.
CC -1- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J03298; AAA0525.1; -;
DR EMBL; D88510; BAA13633.1; -;
DR EMBL; M74778; AAA39427.1; -;
DR PIR; A28438; A28438.
DR HSSP; P02788; 1LPH.
DR MGD; MGI:96837; Ltf.
DR InterPro; IPR001156; Transferrin.
DR Pfam; PF00405; transferrin; 2.
DR PRINTS; PRO0422; TRANSFERRIN.
DR SMART; SM00094; TR_FER; 2.
DR PROSITE; PS00205; TRANSFERRIN_1; 1.
DR PROSITE; PS00206; TRANSFERRIN_2; 2.
DR PROSITE; PS00207; TRANSFERRIN_3; 2.
KW Iron transport; Glycoprotein; Metal-binding; Repeat; Signal.
FT SIGNAL 1 19
FT CHAIN 20 707
FT REPEAT 20 357
FT REPEAT 358 707
FT REPEAT 358 707
FT DISULFID 27 63
FT DISULFID 37 54
FT DISULFID 133 216
FT DISULFID 175 191
FT DISULFID 188 199
FT DISULFID 249 263
FT DISULFID 366 398
FT DISULFID 376 389
FT DISULFID 423 702
FT DISULFID 443 665
FT DISULFID 475 550
FT DISULFID 499 693
FT DISULFID 509 523
FT DISULFID 520 533
FT DISULFID 591 605
FT DISULFID 643 648
FT METAL 78 78
FT METAL 110 110
FT METAL 210 210
FT METAL 271 271
FT METAL 413 413
FT METAL 451 451

FT METAL 544 544
FT METAL 613 613
FT BINDING 139 139
FT BINDING 481 481
FT CARBOHYD 118 118
FT CARBOHYD 494 494
FT CONFLICT 1 2
FT CONFLICT 25 25
FT CONFLICT 82 82
FT CONFLICT 359 359
FT CONFLICT 382 382
FT CONFLICT 449 449
FT CONFLICT 629 629
SQ SEQUENCE 707 AA; 77761 MW; DD232426DF6169EE CRC54;
Query Match 71.7%; Score 43; DB 1; Length 707;
Best Local Similarity 66.7%; Pred. No. 2.4;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 ARRVVWAAGV 12
DB 358 ATRAVTWCAVG 369
ID CAHB_HUMAN STANDARD; PRT; 328 AA.
AC 075493; O60596; Q9UEC4;
DT 30-MAY-2000 (Rel. 39, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CARBONIC ANHYDRASE-RELATED PROTEIN 2 PRECURSOR (CARP 2) (CA-RP II)
DE (CA-XI) (CARBONIC ANHYDRASE-RELATED PROTEIN 1) (CARP XI).
GN CALL OR CARP2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99097242; PubMed=9878543;
RA Bellingham J., Gregory-Evans K., Gregory-Evans C.Y.;
RT "Sequence and tissue expression of a novel human carbonic
RT anhydrase-related protein, CARP-2, mapping to chromosome 19q13.3";
RL Biochem. Biophys. Res. Commun. 253:364-367(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE=Pancreas;
RX MEDLINE=99280326; PubMed=10350627;
RA Fujikawa-Adachi K., Nishimori I., Taguchi T., Yori K., Onishi S.;
RT "cDNA sequence, mRNA expression, and chromosomal localization of human
RT carbonic anhydrase-related protein, CA-RP XI";
RN Biochim. Biophys. Acta 1431:518-524(1999).
RN [3]
RP SEQUENCE OF 48-314 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=99097349; PubMed=9878252;
RA Lovejoy D.A., Hewett-Emmett D., Porter C.A., Cepol D., Sheffield A.,
RA Vale W.W., Tashian R.E.;
RT "Evolutionarily conserved, 'acatalytic' carbonic anhydrase-related
RT protein XI contains a sequence motif present in the neuropeptide
RT sauvagine: the human CA-RP XI gene (CALL) is embedded between the
RT secretor gene cluster and the DBP gene at 19q13.3";
RL Genomics 54:484-493(1998).
CC -1- FUNCTION: DOES NOT HAVE A CATALYTIC ACTIVITY.
CC -1- SUBCELLULAR LOCATION: SECRETED (POTENTIAL).
CC -1- TISSUE SPECIFICITY: EXPRESSED ABUNDANTLY IN THE BRAIN WITH
CC MODERATE EXPRESSION ALSO PRESENT IN SPINAL CORD AND THYROID.
CC -1- SIMILARITY: BELONGS TO THE EUKARYOTIC-TYPE CARBONIC ANHYDRASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; AF067662; AAC99689.1; -
 DR EMBL; AB018195; BAA36840.1; -
 DR EMBL; AF050106; AAD08802.1; -
 DR HSSP; P00318; 1BV3.
 DR MIW; 604644; -
 DR InterPro; IPR001148; Carb_anhydrase.
 DR Pfam; PF00194; Carb_anhydrase; 1.
 DR ProDom; PD000865; Carb_anhydrase; 1.
 DR PROSITE; PS00162; EUK_CO2_ANGHYDRASE; FALSE_NEG.
 KW Signal.
 FT SIGNAL 1 23 POTENTIAL.
 FT CHAIN 24 328 CARBONIC ANHYDRASE-RELATED PROTEIN 2.
 FT CONFLICT 23 24 AH -> GN (IN REF. 2).
 FT CONFLICT 75 75 L -> V (IN REF. 1).
 FT CONFLICT 280 280 I -> M (IN REF. 3).
 SQ SEQUENCE 328 AA; 36238 MW; A4FC408718B2E857 CRC64;

Query Match 66.7%; Score 40; DB 1; Length 328;
 Best Local Similarity 66.7%; Pred. No. 4.1;
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 ARRARVWAAVG 12
 | | | | | | | | | | | | | |
 Db 9 APRALVWAAVG 20

RESULT 10
 C13A_MYCTU
 ID C13A_MYCTU STANDARD; PRT; 449 AA.
 AC O08447;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE PUTATIVE CYTOCHROME P450 135A1 (EC 1.14.-.-).
 GN CYP135A1 OR RV0327C OR MT0342 OR MTCV63.32C.
 OS Mycobacterium tuberculosis.
 CC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 CC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch K., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence."
 RL Nature 393:537-544(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains."
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.

CC -----
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; Z96800; CAB09576.1; -
 DR EMBL; AE006940; AAK44565.1; -
 DR TIGR; MT0342; -
 DR TubercuList; RV0327c; -
 DR InterPro; IPR001128; Cyt_P450.
 DR Pfam; PF00067; p450; 2.
 DR PRINTS; PR00359; BP450.
 DR PRINTS; PR00385; P450.
 DR PRINTS; PR00463; EP450IV.
 DR PRINTS; PR00465; EP450IV.
 DR PROSITE; PS00086; CYTOCHROME_P450; 1.
 KW Hypothetical protein; Oxidoreductase; Monooxygenase; Heme;
 KW Complete proteome.
 FT BINDING 383
 FT SEQUENCE 449 AA; 50010 MW; 192E3CF1855EDCF5 CRC64;

Query Match 66.7%; Score 40; DB 1; Length 449;
 Best Local Similarity 70.0%; Pred. No. 5.5;
 Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 RARVVWAAVG 12
 : : : : : : : : : : : : : :
 Db 285 QSQVWAAVG 294

RESULT 11
 REP2_ECOLI
 ID REP2_ECOLI STANDARD; PRT; 285 AA.
 AC P03066; Q47411;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE REPLICATION INITIATION PROTEIN.
 GN REPA OR REPAL.
 OS Escherichia coli.
 OG Plasmid IncFII R100, Plasmid IncFII R1, and Plasmid IncFII NR1.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC PLASMID=IncFII R100;
 RX MEDLINE=81074309; PubMed=7003300;
 RA Rosen J., Ryder T., Inokuchi H., Ohtsubo H., Ohtsubo E.;
 RT "Genes and sites involved in replication and incompatibility of an
 RT R100 plasmid derivative based on nucleotide sequence analysis."
 RL Mol. Gen. Genet. 179:527-537(1980).
 RN [2]
 RP SEQUENCE OF 1-77 FROM N.A.
 RC PLASMID=IncFII R100, and IncFII R1;
 RX MEDLINE=81173118; PubMed=6163994;
 RA Rosen J., Ryder T., Ohtsubo H., Ohtsubo E.;
 RT "Role of RNA transcripts in replication incompatibility and copy
 RT number control in antibiotic resistance plasmid derivatives."
 RL Nature 290:794-797(1981).
 RN [3]
 RP SEQUENCE OF 270-285 FROM N.A.
 RC PLASMID=IncFII R1;
 RX MEDLINE=86289416; PubMed=3041379;
 RA Masai H., Arai K.;
 RT "RepA protein- and oriR-dependent initiation of R1 plasmid
 RT replication: identification of a rho-dependent transcription

```

RT RL Nucleic Acids Res. 16:6493-6514(1988).
RN [4]
RP SEQUENCE FROM N.A.
RC PLASMID-IncFII NR1;
RX MEDLINE=89011975; PubMed=3050127;
RA Dong X., Womble D.D., Rownd R.H.;
RT "In-vivo studies on the cis-acting replication initiator protein of
IncFII plasmid NR1.";
RL J. Mol. Biol. 202:495-509(1988).
RN [5]
RP SEQUENCE FROM N.A.
RC PLASMID-IncFII NR1;
RX MEDLINE=85160860; PubMed=2580099;
RA Womble D.D., Sampathkumar P., Easton A.M., Luckow V.A., Rownd R.H.;
RT "Transcription of the replication control region of the IncFII
R-plasmid NR1 in vitro and in vivo.";
RL J. Mol. Biol. 181:395-410(1985).
RN [6]
RP SEQUENCE FROM N.A.
RC PLASMID-IncFII R100;
RX MEDLINE=86319522; PubMed=3019092;
RA Ohtsubo H., Ryder T.B., Maeda Y., Armstrong K., Ohtsubo E.;
RT "DNA replication of the resistance plasmid R100 and its control.";
RL Adv. Biophys. 21:115-133(1986).
CC -1- FUNCTION: THIS PROTEIN IS ESSENTIAL FOR PLASMID REPLICATION; IT
IS INVOLVED IN COPY CONTROL FUNCTIONS.
CC -1- SIMILARITY: BELONGS TO THE INCFII REPA FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; J01762; AAA92257.1; -
CC EMBL; J01770; -; NOT_ANNOTATED_CDS.
CC EMBL; X12587; CAA31100.1; -
CC EMBL; X12776; CAA31263.1; -
CC EMBL; X02302; CAA26168.1; -
CC EMBL; X02302; CAA26169.1; ALT_INIT.
CC EMBL; M26840; AAA26067.1; -
CC PIR; A03602; IDECRP.
CC PIR; S01773; S01773.
CC InterPro; IPR003446; IncFII_repA.
CC Pfam; PF02387; IncFII_repA; 1.
KW Plasmid; DNA replication; Plasmid copy control.
FT CONFLICT 55 55 R -> H (IN REF. 6).
SQ SEQUENCE 285 AA; 32755 MW; A21C9D59D24B26BB CRC64;

Query Match 65.0%; Score 39; DB 1; Length 285;
Best Local Similarity 87.5%; Pred. No. 5.4;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ARRRVW 8
Db 168 ARRSRVW 175

RESULT 12
YH92_AERPE
ID YH92_AERPE STANDARD; PRT; 176 AA.
AC Q9YB03;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HYPOTHETICAL PROTEIN APEI1792.
GN Apep1792.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Desulfurococcaceae; Desulfurococcaceae;

terminator required for cis-action of repA protein.";
RN [4]
RP SEQUENCE FROM N.A.
RC PLASMID-IncFII NR1;
RX MEDLINE=89011975; PubMed=3050127;
RA Dong X., Womble D.D., Rownd R.H.;
RT "In-vivo studies on the cis-acting replication initiator protein of
IncFII plasmid NR1.";
RL J. Mol. Biol. 202:495-509(1988).
RN [5]
RP SEQUENCE FROM N.A.
RC PLASMID-IncFII NR1;
RX MEDLINE=85160860; PubMed=2580099;
RA Womble D.D., Sampathkumar P., Easton A.M., Luckow V.A., Rownd R.H.;
RT "Transcription of the replication control region of the IncFII
R-plasmid NR1 in vitro and in vivo.";
RL J. Mol. Biol. 181:395-410(1985).
RN [6]
RP SEQUENCE FROM N.A.
RC PLASMID-IncFII R100;
RX MEDLINE=86319522; PubMed=3019092;
RA Ohtsubo H., Ryder T.B., Maeda Y., Armstrong K., Ohtsubo E.;
RT "DNA replication of the resistance plasmid R100 and its control.";
RL Adv. Biophys. 21:115-133(1986).
CC -1- FUNCTION: THIS PROTEIN IS ESSENTIAL FOR PLASMID REPLICATION; IT
IS INVOLVED IN COPY CONTROL FUNCTIONS.
CC -1- SIMILARITY: BELONGS TO THE INCFII REPA FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; J01762; AAA92257.1; -
CC EMBL; J01770; -; NOT_ANNOTATED_CDS.
CC EMBL; X12587; CAA31100.1; -
CC EMBL; X12776; CAA31263.1; -
CC EMBL; X02302; CAA26168.1; -
CC EMBL; X02302; CAA26169.1; ALT_INIT.
CC EMBL; M26840; AAA26067.1; -
CC PIR; A03602; IDECRP.
CC PIR; S01773; S01773.
CC InterPro; IPR003446; IncFII_repA.
CC Pfam; PF02387; IncFII_repA; 1.
KW Plasmid; DNA replication; Plasmid copy control.
FT CONFLICT 55 55 R -> H (IN REF. 6).
SQ SEQUENCE 285 AA; 32755 MW; A21C9D59D24B26BB CRC64;

Query Match 63.3%; Score 38; DB 1; Length 176;
Best Local Similarity 70.0%; Pred. No. 5.3;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 RARVWVAVG 12
Db 72 RPRVWVGVG 81

RESULT 13
GALT_HUMAN
ID GALT_HUMAN STANDARD; PRT; 368 AA.
AC G0755;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE GALANIN RECEPTOR TYPE 3 (GAL3-R) (GALR3).
GN GALR3 OR GALNR3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98389766; PubMed=9722565;
RA Smith K.E., Walker M.W., Artymyshyn R., Bard J., Borowsky B.,
RA Tamm J.A., Yao W.-J., Vaysse P.J.-J., Branchek T.A., Gerald C.,
RA Jones K.A.;
RT "Cloned human and rat galanin GALR3 receptors: pharmacology and
activation of G-protein inwardly rectifying K+ channels.";
RL J. Biol. Chem. 273:23321-23326(1998).
RN [2]
RP SEQUENCE FROM N.A.
TI Tissue-Uterus;
RC Bennett M.M., Lescoe M.K., Gallipoli P.Z., Ramabhadran T.V.;
RT "Homologue of the human galanin 2 receptor gene isolated from a human
uterus cDNA library.";
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=99048961; PubMed=9832121;
RA Kolakowski L.F. Jr., O'Neill G.P., Howard A.D., Broussard S.R.,
RA Sullivan K.A., Feighner S.D., Sawdzargo M., Nguyen T., Kargman S.,

```

RA Shiao L.-L., Hreniuk D.L., Tan C.P., Evans J., Abramowitz M.,
RA Chateaufneuf A., Coulombe N., Ng G., Johnson M.P., Tharian A.,
RA Khoshbouei H., George S.R., Smith R.G., O'Dowd B.F.,
RT "Molecular characterization and expression of cloned human galanin
RT receptors GALR2 and GALR3".
RL J. Neurochem. 71:2239-2251(1998).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE-99127120; PubMed-9928159;
RA Iismaa T.P., Fathi Z., Hort Y.J., Iben L.G., Dutton J.L., Baker E.,
RA Sutherland G.R., Shine J.,
RT "Structural organization and chromosomal localization of three human
RT galanin receptor genes".
RL Ann. N.Y. Acad. Sci. 863:56-63(1998).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE-20057165; PubMed-10591208;
RA Dunham I., Hunt A.R., Collins J.E., Bruskiewich R., Beare D.M.,
RA Clamp M., Smink L.J., Alnsough R., Almeida J.P., Babbage A.,
RA Baguley C., Bailey J., Barlow K., Bates K.N., Beasley O., Bird C.P.,
RA Blakey S., Bridgeman A.M., Buck D., Burgess J., Burrill W.D.,
RA Burton J., Carter C., Carter N.P., Chen Y., Clark G., Clegg S.M.,
RA Cobley V., Cole C.G., Collier R.E., Connor R.E., Conroy D., Corby N.,
RA Coville G.J., Cox A.V., Davis J., Dawson E., Dhami P.D., Dockree C.,
RA Dodsworth S.J., Durbin R.M., Ellington A., Evans K.L., Fey J.M.,
RA Fleming K. French L., Garner A.A., Gilbert J.G.R., Goward M.E.,
RA Grahame D., Griffiths M.N., Hall C., Hall R., Hall-Tamlyn G.,
RA Heathcote R.W., Ho S., Holmes S., Hunt S.E., Jones M.C., Kershaw J.,
RA Kimberley A., King A., Laird G.D., Langford C.F., Leversha M.A.,
RA Lloyd C., Lloyd D.M., Martyn I.K., Mashreghi-Mohammadi M.,
RA Matthews L., McCann O.T., McClay J., McLaren S., McMurray A.A.,
RA Milne S.A., Mortimore B.J., Odeil C.N., Pavitt R., Pearce A.V.,
RA Pearson D., Phillips M.J., Phillips S.H., Plumb R.W., Ramsay H.,
RA Ramsey Y., Rogers L., Ross M.T., Scott C.E., Sehra H.K., Skuce C.D.,
RA Smalley S., Smith M.L., Soderlund C., Spragon L., Steward C.A.,
RA Sulston J.E., Swann R.M., Vaudin M., Wall M., Wallis J.M.,
RA Whiteley M.N., Willey D., Williams L., Williams S., Williamson D.,
RA Wilmer T.E., Wilming L., Wright C.L., Hubbard T., Bentley D.R.,
RA Beck S., Rogers J., Shimizu N., Minoshima S., Kawasaki K., Sasaki T.,
RA Asakawa S., Kudoh J., Shintani A., Shibuya K., Yoshizaki Y., Aoki N.,
RA Mitsuyama S., Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S.,
RA Do A., Do T., Dorman A., Fang F., Fu Y., Hu P., Hua A., Kenton S.,
RA Lai H., Lao H.L., Lewis J., Lewis S., Lin S.-P., Loh P., Mala E.,
RA Nguyen T., Pan H., Phan S., Qi S., Qian Y., Ray L., Ren Q., Shaull S.,
RA Sloan D., Song L., Wang Q., Wang Y., Wang Z., White J., Willingham D.,
RA Wu H., Yao Z., Zhan M., Zhang G., Chissos S., Murray J., Miller N.,
RA Minx P., Fulton R., Johnson D., Bemis G., Bentley D., Bradshaw H.,
RA Bourne S., Cordes M., Du Z., Fulton L., Goela D., Graves T.,
RA Hawkins J., Hinds K., Kemp K., Latreille P., Layman D., Ozersky P.,
RA Rohlfing T., Scheet P., Walker C., Wamsley A., Wohlmann P., Pepin K.,
RA Nelson J., Korf I., Bedell J.A., Hillier L., Mardis E., Waterston R.,
RA Wilson R., Emanuel B.S., Shaikh T., Kurahashi H., Saitta S.,
RA Budarf M.L., McDermid H.E., Johnson A., Wong A.C.C., Morrow B.E.,
RA Edelmann L., Kim U.J., Shizuya H., Simon M.I., Dumanski J.P.,
RA Peyrard M., Kedra D., Seroussi E., Franconi I., Tapia I., Bruder C.E.,
RA O'Brien K.P., Wilkinson P., Bodenteich A., Hartman K., Hu X.,
RA Khan A.S., Lane L., Tilahun Y., Wright H.,
RT "The DNA sequence of human chromosome 22".
RL Nature 402:489-495(1999).
CC -1- FUNCTION: RECEPTOR FOR THE HORMONE GALANIN.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF073799; AAC35944.1; -.
DR EMBL; AF067733; AAC18860.1; -.

DR EMBL; AF129514; AADA7348.2; -.
DR EMBL; AF129513; AADA7348.2; JOINED.
DR EMBL; Z97630; CAB42831.1; -.
DR MIM; 603692; -.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR InterPro; IPR003908; Galnln3_rcptor.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHHODPSN.
DR PRINTS; PR00663; GALANINR.
DR PROSITE; PS00237; G-PROTEIN_RECP_F1_1; 1.
DR PROSITE; PS0262; G-PROTEIN_RECP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
Multigene family; Lipoprotein; Palmitate.
FT DOMAIN 1 20 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 21 41 1 (POTENTIAL).
FT DOMAIN 42 57 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 58 78 2 (POTENTIAL).
FT DOMAIN 79 96 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 97 118 3 (POTENTIAL).
FT DOMAIN 119 138 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 139 159 4 (POTENTIAL).
FT DOMAIN 160 184 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 185 205 5 (POTENTIAL).
FT DOMAIN 206 236 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 237 257 6 (POTENTIAL).
FT DOMAIN 258 259 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 260 280 7 (POTENTIAL).
FT DOMAIN 281 368 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 6 6 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 95 172 BY SIMILARITY.
FT LIPID 308 308 PALMITATE (BY SIMILARITY).
FT DOMAIN 218 225 POLY-ALA.
FT DOMAIN 310 318 POLY-ARG.
SQ SEQUENCE 368 AA; 39573 MW; E8F46866422C3150 CRC64;

Query Match 61.7%; Score 37; DB 1; Length 368;
Best Local Similarity 63.6%; Pred. No. 16;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 RRARVWNAVG 12

Db 206 RFLRFLNAVG 216

RESULT 14

BGLR_HUMAN STANDARD; PRT; 651 AA.
AC P08236;

DT 01-AUG-1988 (Rel. 08, Created)

DT 01-AUG-1988 (Rel. 08, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last annotation update)

DE BETA-GLUCURONIDASE PRECURSOR (EC 3.2.1.31) (BETA-G1).

GN GUSB.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

{}]

RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.

RC TISSUE=Placenta;

RX MEDLINE-87118233; PubMed-3468507;

RA Oshima A., Kyle J.W., Miller R.D., Hoffmann J.W., Powell P.P.,

RA Grubb J.H., Sly W.S., Tropak M., Guise K.S., Gravel R.A.;

RT "Cloning, sequencing, and expression of cDNA for human beta-

RT glucuronidase".

RL Proc. Natl. Acad. Sci. U.S.A. 84:685-689(1987).

RN [2]

RP SEQUENCE OF 1-70 FROM N.A.

RX MEDLINE-92009900; PubMed-1916806;

RA Shipley J.M., Miller R.D., Wu B.M., Grubb J.H., Christensen S.G.,

RA Kyle J.W., Sly W.S.;

RT "Analysis of the 5' flanking region of the human beta-glucuronidase

gene.";
Genomics 10:1009-1018(1991).
[3]
SEQUENCE OF 520-585 FROM N.A.
RC TISSUE-Fibroblast;
RX MEDLINE-85232043; PubMed-3924735;
RA Guise K.S., Korneluk R.G., Wayne J., Lamhonwah A.-M., Quan F.,
Palmer R., Ganschow R.E., Sly W.S., Gravel R.A.;
"Isolation and expression in *Escherichia coli* of a cDNA clone encoding
human beta-glucuronidase.";
RL Gene 34:105-110(1985).
[4]
SEQUENCE OF 23-32 AND 160-175.
RC TISSUE-Placenta;
RX MEDLINE-92162201; PubMed-1311180;
RA Tanaka J., Gasa S., Sakurada K., Miyazaki T., Kasai M., Makita A.;
"Characterization of the subunits and sugar moiety of human placental
and leukemic beta-glucuronidase.";
RL Biol. Chem. Hoppe-Seyler 373:57-62(1992).
[5]
X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).
RX MEDLINE-96185449; PubMed-8593764;
RA Jain S., Drendel W.B., Chen Z.W., Mathews F.S., Sly W.S., Grubb J.H.;
"Structure of human beta-glucuronidase reveals candidate lysosomal
targeting and active-site motifs.";
RL Nat. Struct. Biol. 3:375-381(1996).
[6]
VARIANT MPS-VII TRP-216.
RX MEDLINE-94154730; PubMed-8111412;
RA Vervoort R., Lissens W., Liebaers I.;
"Molecular analysis of a patient with hydrops fetalis caused by beta-
glucuronidase deficiency, and evidence for additional pseudogenes.";
RL Hum. Mutat. 2:443-445(1993).
[7]
VARIANTS MPS-VII VAL-354 AND TRP-611.
RX MEDLINE-94154731; PubMed-8111413;
RA Wu B.M., Sly W.S.;
"Mutational studies in a patient with the hydrops fetalis form of
mucopolysaccharidosis type VII.";
RL Hum. Mutat. 2:446-457(1993).
[8]
VARIANTS MPS-VII CYS-382 AND VAL-619.
RX MEDLINE-91090114; PubMed-1702266;
RA Tomatsu S., Fukuda S., Sukeyama K., Ikeda Y., Yamada S., Yamada Y.,
Sasaki T., Okamoto H., Kuwahara T., Yanaguchi S., Kinan T.,
Shintaku H., Ishiki G., Orii T.;
"Mucopolysaccharidosis type VII: characterization of mutations and
molecular heterogeneity.";
RL Am. J. Hum. Genet. 48:89-96(1991).
[9]
VARIANT MPS-VII CYS-627.
RX MEDLINE-93190983; PubMed-7680524;
RA Shipley J.M., Klinkenberg M., Wu B.M., Bachinsky D.R., Grubb J.H.,
Sly W.S.;
"Mutational analysis of a patient with mucopolysaccharidosis type
VII, and identification of pseudogenes.";
RL Am. J. Hum. Genet. 52:517-526(1993).
CC -1- FUNCTION: PLAYS AN IMPORTANT ROLE IN THE DEGRADATION OF DERMATAN
AND KERATAN SULFATES.
CC -1- CATALYTIC ACTIVITY: A BETA-D-GLUCURONOSIDE + H(2)O -> AN
ALCOHOL + D-GLUCURONATE.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SUBCELLULAR LOCATION: LYSOSOMAL.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- PTM: GLYCOSYLATED WITH 3 TO 4 N-LINKED OLIGOSACCHARIDE CHAINS.
CC -1- DISEASE: DEFECTS IN GUSB ARE THE CAUSE OF MUCOPOLYSACCHARIDOSIS
TYPE VII (MPS-VII) (ALSO KNOWN AS SLY SYNDROME). THIS IS AN
AUTOSOMAL RECESSIVE DISORDER.
CC -1- SIMILARITY: BELONGS TO FAMILY 2 OF GLYCOSYL HYDROLASES.

THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL: M15182; AAA52561.1; -
CC EMBL: M65002; AAA52622.1; -
CC EMBL: M10618; AAA52621.1; -
CC EMBL: S72462; AAD14101.1; -
CC PIR: A26581; A26581.
CC PDB: 1BHG; 17-SEP-97.
CC MIM: 253220; -
CC InterPro: IPR001649; Glyco_hydro_2.
CC Pfam: PF00703; Glyco_hydro_2; 1.
CC PRINTS: PR00132; GLHYDRLASE2.
CC PROSITE: PS00719; GLYCOSYL_HYDROL_F2_1; 1.
CC PROSITE: PS00608; GLYCOSYL_HYDROL_F2_2; 1.
CC KW Hydrolase; Glycosidase; Lysosome; Glycoprotein; Signal;
KW Mucopolysaccharidosis; Disease mutation; 3D-structure;
KW Alternative splicing.
FT SIGNAL 1 22 BETA-GLUCURONIDASE.
FT CHAIN 23 651 PROTON DONOR.
FT ACT_SITE 451 451 N-LINKED (GLCNAC. . .).
FT CARBOHYD 173 173 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 272 272 N-LINKED (GLCNAC. . .).
FT CARBOHYD 420 420 N-LINKED (GLCNAC. . .).
FT CARBOHYD 631 631 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 305 355 MISSING (IN SHORT ISOFORM).
FT VARIANT 216 216 R -> W (IN MPS-VII).
FT VARIANT 354 354 /FTID-VAR_003196.
FT VARIANT 382 382 A -> V (IN MPS-VII).
FT VARIANT 611 611 R -> C (IN MPS-VII).
FT VARIANT 619 619 R -> W (IN MPS-VII).
FT VARIANT 619 619 /FTID-VAR_003199.
FT VARIANT 619 619 A -> V (IN MPS-VII).
FT VARIANT 627 627 /FTID-VAR_003200.
FT VARIANT 627 627 W -> C (IN MPS-VII).
FT VARIANT 627 627 /FTID-VAR_003201.
SQ SEQUENCE 651 AA; 74715 MW; 6BA7A19935C9ABBD CRC64;

Query Match 61.7%; Score 37; DB 1; Length 651;
Best Local Similarity 58.3%; Pred. No. 27;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ARRARYVVAAGV 12
|| : | ||| : |
DB 2 ARGSAVAWAALG 13

RESULT 15
CAHE SHEEP
ID CAHB_SHEEP STANDARD; PRT; 328 AA.
AC Q95203;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE CARBONIC ANHYDRASE-RELATED PROTEIN 2 PRECURSOR (CARP 2) (CA-RP II)
DE (CA-XI).
GN CALL OR CARP2.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID:9940;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Brain;
RX MEDLINE-99097349; PubMed-9878252;
RA Lovejoy D.A., Hewett-Emmett D., Porter C.A., Cepoi D., Sheffield A.,

RA Vale W.W., Tashian R.E.;
RT "Evolutionarily conserved, 'acatalytic' carbonic anhydrase-related
RT protein XI contains a sequence motif present in the neuropeptide
RT sauvagine: the human CA-Rp XI gene (CAll) is embedded between the
RT secretor gene cluster and the DBP gene at 19q13.3";
RL Genomics 54:484-493(1998).
CC -|- FUNCTION: DOES NOT HAVE A CATALYTIC ACTIVITY.
CC -|- SUBCELLULAR LOCATION: SECRETED (POTENTIAL).
CC -|- SIMILARITY: BELONGS TO THE EUKARYOTIC-TYPE CARBONIC ANHYDRASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Y07785; CAA69107.1; -.
DR HSSP; P00918; 1BV3.
DR InterPro; IPR001148; Carb_anhydrase.
DR Pfam; PF00194; carb_anhydrase; 1.
DR ProDom; PD000865; Carb_anhydrase; 1.
DR PROSITE; PS00162; EUK_CO2_ANHYDRASE; FALSE_NEG.
KW Signal.
FT SIGNAL 1 23 POTENTIAL.
FT CHAIN 24 328 CARBONIC ANHYDRASE-RELATED PROTEIN 2.
SQ SEQUENCE 328 AA; 36199 MW; E28A6C7236EC41C3 CRC64;

Query Match 60.0%; Score 36; DB 1; Length 328;
Best Local Similarity 58.3%; Pred. No. 21;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ARARVVWAAG 12
| | | | |
Db 9 APRVLVWALG 20

Search completed: December 31, 2001, 08:53:05
Job time: 59 sec

1	56	93.3	711	4	Q9H123
2	43	71.7	491	2	Q59702
3	40	66.7	349	5	Q9VB34
4	40	66.7	504	13	Q98923
5	39	65.0	146	11	Q9DAB9
6	38	63.3	18	6	Q9TR68
7	38	63.3	342	10	Q82517
8	38	63.3	431	2	Q9L0G9
9	38	63.3	509	10	Q41755
10	38	63.3	567	10	Q65341
11	38	63.3	567	10	Q65342
12	38	63.3	655	10	Q9LKI1
13	38	63.3	986	2	Q9A457
14	37	61.7	108	1	Q9YA23
15	37	61.7	256	2	Q9KZ43
16	37	61.7	338	2	Q31936
17	37	61.7	338	9	Q64065
18	37	61.7	378	2	Q9RK67
19	37	61.7	397	2	Q9RWD4
					Q9H123 homo sapien
					Q59702 pseudomonas
					Q9VB34 drosophila
					Q98923 gallus galli
					Q9DAB9 mus musculus
					Q9TR68 bos taurus
					Q82517 saccharum h
					Q9L0G9 streptomyces
					Q41755 zea mays (m
					Q65341 saccharum r
					Q65342 saccharum o
					Q9LKI1 oryza sativ
					Q9A457 caulobacter
					Q9YA23 aeropyrum p
					Q9KZ43 streptomyces
					Q31936 bacillus su
					Q64065 bacterioph
					Q9RK67 streptomyces
					Q9RWD4 deinococcus

Q959702;
 01-NOV-1996 (Tremblrel. 01, Created)
 01-NOV-1999 (Tremblrel. 12, Last sequence update)
 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 P-HYDROXYBENZALDEHYDE DEHYDROGENASE.
 PCBA.
 OS Pseudomonas putida.
 OG Plasmid pRA000.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas
 OC NCBI_TaxID=303;
 RN [1]
 RN PARTIAL SEQUENCE FROM N.A.
 RC STRAIN=NCIMB 9866;
 RC MEDLINE=95014081; PubMed=7929007;
 RA Kim J., Fuller J.H., Cecchini G., McIntire W.S.;
 RA "Cloning, sequencing, and expression of the structural genes for the
 RT cytochrome and flavoprotein subunits of p-cresol methylhydroxylase
 RT from two strains of Pseudomonas putida.";
 RN J. Bacteriol. 176:6349-6361(1994).
 RN [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=NCIMB 9866;
 RC MEDLINE=20029269; PubMed=10565539;
 RA Cronin C.N., Kim J.-H., Fuller J., Zhang X.-P., McIntire W.S.;
 RA "Organization and sequences of p-hydroxybenzaldehyde dehydrogenase and
 RT other plasmid-encoded genes for early enzymes of the p-cresol
 RT degradative pathway in Pseudomonas putida NCIMB 9866 and 9869.";
 RN DNA Seq. 10:7-17(1999).
 RN [3]
 RN SEQUENCE FROM N.A.
 RC STRAIN=NCIMB 9866;
 RC STRAIN=NCIMB 9866;
 RA Cronin C.N.;
 RA Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
 RL EMBL; U96338; AAA75634.2;
 DR HSSP; P56533; 1A4S.
 DR InterPro; IPR002086; Aldehyde_dehydr.
 DR Pfam; PF00171; aldehyd; 1.
 DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
 DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
 KW Plasmid.
 SQ SEQUENCE 491 AA; 53156 MW; 70C45CEA21266344 CRC64;
 Query Match 71.7%; Score 43; DB 2; Length 491;
 Best Local Similarity 66.7%; Pred. No. 14;
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 ARARVWVAAG 12
 Db 58 ARQAQVANAAG 69
 ||:|:|:|:|
 RESULT 3
 Q9VB34 PRELIMINARY; PRT; 349 AA.
 ID Q9VB34
 AC Q9VB34;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE CG13973 PROTEIN.
 GN SIDE OR CG12510 OR CG13973 OR CG13974 OR CG13975 OR CG13979.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OC NCBI_TaxID=7227;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RA MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,

George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 Sutton G.G., Wortman J.R., Yeaman M.D., Zhang Q., Chen L.X.,
 Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 Burks K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
 Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 Hosin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
 Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 Palazzolo M., Pittman G.S., Pan S., Poilard J., Puri V., Reese M.G.,
 Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J., Yao Q.A.,
 Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Zhao Q., Zheng L.,
 Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhou X., Zhu S., Zhu X., Smith H.O.,
 Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 "The genome sequence of Drosophila melanogaster.";
 Science 287:2185-2195(2000).
 CC -1- SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
 DOMAIN.
 DR EMBL; AE003760; AAF56709.1;
 DR FlyBase; FBgn0016061; side.
 DR InterPro; IPR003598; Ig_c2.
 DR InterPro; IPR003600; Ig_Like.
 DR InterPro; IPR003006; Ig_MHC.
 DR Pfam; PF00047; ig; 3.
 DR SMART; SM00408; IGC2; 1.
 DR SMART; SM00410; IG_Like; 1.
 SQ SEQUENCE 349 AA; 39000 MW; 4F3877C6DC22CC5C CRC64;
 Query Match 66.7%; Score 40; DB 5; Length 349;
 Best Local Similarity 77.8%; Pred. No. 32;
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 Qy 4 ARVWVAAG 12
 Db 9 ARLVWRAAG 17
 ||:|:|:|:|
 RESULT 4
 Q98923 PRELIMINARY; PRT; 504 AA.
 ID Q98923
 AC Q98923;
 DT 01-FEB-1997 (Tremblrel. 02, Created)
 DT 01-FEB-1997 (Tremblrel. 02, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE HEMCAM PRECURSOR.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OC NCBI_TaxID=9031;
 RN [1]
 RN SEQUENCE FROM N.A.

RC STRAIN-H.B13OV4; TISSUE=THYMUS;
 RX MEDLINE-97133433; PubMed-8978830;
 RA Vainio O., Dunon D., Aissi F., Dangy J.P., McNaghy K.M., Imhof B.A.;
 RT "HEMCAm, an adhesion molecule expressed by c-kit+ hemopoietic
 progenitors.";
 RL J. Cell Biol. 135:1655-1668(1996).
 CC -1- SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
 CC DOMAIN.
 DR EMBL; Y08854; CAA70079.1; -;
 DR InterPro; IPR003598; Ig_C2.
 DR InterPro; IPR003600; Ig-like.
 DR InterPro; IPR003006; Ig_MHC.
 DR Pfam; PF00047; Ig; 5.
 DR SMART; SM00408; Igc2; 1.
 DR SMART; SM00410; Ig-like; 2.
 KW Signal.
 FT SIGNAL 1 35 POTENTIAL.
 FT CHAIN 36 504 POTENTIAL.
 SQ SEQUENCE 504 AA; 55540 MW; ADAE4B94ED4F02E1 CRC64;
 Query Match 56.7%; Score 40; DB 13; Length 504;
 Best Local Similarity 63.6%; Pred. No. 47;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 2 RRARVWAAVG 12
 Db 487 RAGRVLWAAAG 497
 RESULT 5
 ID Q9DAB9 PRELIMINARY; PRT; 146 AA.
 AC Q9DAB9;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE 1700014B07RIK PROTEIN.
 GN 1700014B07RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-C57BL/6J; TISSUE=TESTIS;
 RX MEDLINE-21085660; PubMed-11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
 RA Schirni L.M., Stauffli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki K., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 DR EMBL; AK005972; BAB24347.1; -;
 DR MGD; MGI:1919498; 1700014B07RIK.
 SQ SEQUENCE 146 AA; 15492 MW; 60B913081A1824CF CRC64;

Query Match 65.0%; Score 39; DB 11; Length 146;

Best Local Similarity 66.7%; Pred. No. 20;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 1 ARRARVWAAVG 12
 Db 59 SRRARARWAGG 70
 RESULT 6
 ID Q9TRG8 PRELIMINARY; PRT; 18 AA.
 AC Q9TRG8;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
 DE LACTOFERRIN (FRAGMENT).
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE
 RX MEDLINE-93253156; PubMed-8486845;
 RA Shimazaki K., Tanaka T., Kon H., Oota K., Kawaguchi A., Maki Y.,
 RA Sato T.;
 RL J. Dairy Sci. 76:946-955(1993).
 SQ SEQUENCE 18 AA; 2138 MW; 809196895D52D7CE CRC64;

Query Match 63.3%; Score 38; DB 6; Length 18;
 Best Local Similarity 87.5%; Pred. No. 3.6;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 RVVWAAVG 12
 Db 3 RVVWAAVG 10

RESULT 7
 ID O82517 PRELIMINARY; PRT; 342 AA.
 AC O82517;
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE SOLUBLE ACID INVERTASE (FRAGMENT).
 GN SCINVH3.2.
 OS Saccharum hybrid cultivar H65-7052.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
 OC Panicoideae; Andropogoneae; Saccharum.
 OX NCBI_TaxID=50503;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. H65-7052;
 RA Albert H.H., Zhu Y.J., Moore P.H.;
 RT "Differential expression of soluble acid invertase (SAI) genes
 correlates to differences in sucrose accumulation in sugarcane.";
 RL Submitted (Aug-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF TERMINAL NON-REDUCING BETA-D-
 CC FRUCTOFURANOSIDE RESIDUES IN BETA-D-FRUCTOFURANOSIDES.
 CC -1- SIMILARITY: BELONGS TO FAMILY 32 OF GLYCOSYL HYDROLASES.
 DR EMBL; AF083856; AAC36118.1; -;
 DR Mendel; 33105; Sacyh; 1002; 33105.
 DR InterPro; IPR001362; Glyco_hydro_32.
 DR Pfam; PF00251; Glyco_hydro_32; 1.
 KW Glycoprotein; Glycosidase; Hydrolase.
 FT NON_TER 1
 SQ SEQUENCE 342 AA; 37279 MW; F073F49B22083544 CRC64;

Query Match 63.3%; Score 38; DB 10; Length 342;

Best Local Similarity 58.3%; Pred. No. 71;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ARARVVWAAVG 12
|:| ||:| ||
Db 84 AKRRVLWGWWG 95

RESULT 8
Q9L0G9 PRELIMINARY; PRT; 431 AA.
AC Q9L0G9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE PUTATIVE NTEGRAL MEMBRANE PROTEIN.
GN SCD63.20.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA Oliver K., Harris D.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kieser H.M., Denapaita D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
DR EMBL; ALJ61755; CAB82027.1; -.
SQ SEQUENCE 431 AA; 44610 MW; 049C7148CF41A5E5 CRC64;

Query Match 63.3%; Score 38; DB 2; Length 431;
Best Local Similarity 80.0%; Pred. No. 90;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 RRARVVWAAV 11
|:| |||||
Db 102 RTARVVWAAV 111

RESULT 9
Q41755 PRELIMINARY; PRT; 509 AA.
ID Q41755;
AC Q41755;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE INVERTASE (FRAGMENT).
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MERIT;
RA Xu J., Almira E.C., Avigne W.T., McCarty D.R., Koch K.E.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF TERMINAL NON-REDUCING BETA-D-
FRUCTOFURANOSIDE RESIDUES IN BETA-D-FRUCTOFURANOSIDES.

-1- SIMILARITY: BELONGS TO FAMILY 32 OF GLYCOSYL HYDROLASES.
CC EMBL; U31451; AAA74584.1; -.
DR Mendel; 11124; Zea; 1002; 11124.
DR InterPro; IPR001362; Glyco_hydro_32.
DR Pfam; PF00251; Glyco_hydro_32; 1.
KW Glycoprotein; Glycosidase; Hydrolase.
FT NON_TER 1
SQ SEQUENCE 509 AA; 55678 MW; 4A74439FDBD8C34E CRC64;

Query Match 63.3%; Score 38; DB 10; Length 509;
Best Local Similarity 58.3%; Pred. No. 1.1e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ARARVVWAAVG 12
|:| ||:| ||
Db 248 AKRRVLWGWWG 259

RESULT 10
O65341 PRELIMINARY; PRT; 567 AA.
ID O65341;
AC O65341;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE SOLUBLE ACID INVERTASE (EC 3.2.1.26) (FRAGMENT).
OS Saccharum robustum.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Saccharum.
OX NCBI_TaxID=62334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. MOLOKAI 5829; TISSUE=STEM APEX;
RA Albert H.H., Zhu Y.J., Moore P.H.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF TERMINAL NON-REDUCING BETA-D-
FRUCTOFURANOSIDE RESIDUES IN BETA-D-FRUCTOFURANOSIDES.
CC -1- SIMILARITY: BELONGS TO FAMILY 32 OF GLYCOSYL HYDROLASES.
DR EMBL; AF062734; AAC16654.1; -.
DR Mendel; 29707; Sacro; 1002; 29707.
DR InterPro; IPR001362; Glyco_hydro_32.
DR Pfam; PF00251; Glyco_hydro_32; 1.
DR PROSITE; PS00609; GLYCOSYL_HYDROL_F32; 1.
KW Glycoprotein; Glycosidase; Hydrolase.
FT NON_TER 1
SQ SEQUENCE 567 AA; 62631 MW; 92FF2DD18772FBB3 CRC64;

Query Match 63.3%; Score 38; DB 10; Length 567;
Best Local Similarity 58.3%; Pred. No. 1.2e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ARARVVWAAVG 12
|:| ||:| ||
Db 309 AKRRVLWGWWG 320

RESULT 11
O65342 PRELIMINARY; PRT; 567 AA.
ID O65342;
AC O65342;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE SOLUBLE ACID INVERTASE (EC 3.2.1.26) (FRAGMENT).
OS Saccharum officinarum (Sugarcane).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Saccharum.
OX NCBI_TaxID=4547;
RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN-CV. LOUISIANA PURPLE; TISSUE-STEM APEX;
 RA Albert H.H., Zhu Y.J., Moore P.H.;
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF TERMINAL NON-REDUCING BETA-D-
 FRUCTOFURANOSIDE RESIDUES IN BETA-D-FRUCTOFURANOSIDES.
 CC -1- SIMILARITY: BELONGS TO FAMILY 32 OF GLYCOSYL HYDROLASES.
 DR EMBL; AF062735; AAC16655.1; -;
 DR InterPro; IPR001362; Glyco_hydro.32.
 DR Pfam; PF00251; Glyco_hydro.32; 1.
 DR PROSITE; PS00609; GLYCOSYL_HYDROL_F32; 1.
 KW Glycoprotein; Glycosidase; Hydrolase.
 FT NON_TER 1
 SQ SEQUENCE 567 AA; 62637 MW; 47A1FC84CFEB2AFB CRC64;

Query Match 63.3%; Score 38; DB 10; Length 567;
 Best Local Similarity 58.3%; Pred. No. 1.2e+02;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ARRARRVWAAVG 12
 I-| | | | |
 Db 309 AKRRRVWGVWG 320

RESULT 12
 Q9LKI9 PRELIMINARY; PRT; 655 AA.
 AC Q9LKI9
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE VACUOLAR ACID INVERTASE (EC 3.2.1.26).
 GN INV3.
 OS Oryza sativa (Rice).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 CC Ehrhartoideae; Oryzeae; Oryza.
 OC NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Fernandes S.O., Dennis E.S., Dolferus R.;
 RT "Cloning of rice invertase genes";
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF TERMINAL NON-REDUCING BETA-D-
 FRUCTOFURANOSIDE RESIDUES IN BETA-D-FRUCTOFURANOSIDES.
 CC -1- SIMILARITY: BELONGS TO FAMILY 32 OF GLYCOSYL HYDROLASES.
 DR EMBL; AF276704; AAF87246.1; -;
 DR InterPro; IPR001362; Glyco_hydro.32.
 DR Pfam; PF00251; Glyco_hydro.32; 1.
 DR PROSITE; PS00609; GLYCOSYL_HYDROL_F32; 1.
 KW Glycoprotein; Glycosidase; Hydrolase.
 SQ SEQUENCE 655 AA; 71765 MW; B5389AB9A1B6AB84 CRC64;

Query Match 63.3%; Score 38; DB 10; Length 655;
 Best Local Similarity 58.3%; Pred. No. 1.4e+02;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ARRARRVWAAVG 12
 I-| | | | |
 Db 405 AKRRRVWGVWG 416

RESULT 13
 Q9A4S7 PRELIMINARY; PRT; 986 AA.
 AC Q9A4S7
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE GLUTAMATE-AMMONIA-LIGASE ADENYLYLTRANSFERASE, PUTATIVE.

GN CC2753.
 OS Caulobacter crescentus.
 CC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
 OC Caulobacter.
 OX NCBI_TaxID=69394;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21173698; PubMed=11259647;
 RA Niernan W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
 RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
 RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
 RA Deboy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
 RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
 RA Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
 RT "Complete genome sequence of Caulobacter crescentus";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
 DR EMBL; AE005941; AAK24717.1; -;
 DR TIGR; CC2753; -;
 KW Transferase; Nucleotidyltransferase; Ligase; Complete proteome.
 SQ SEQUENCE 986 AA; 106530 MW; 8B4B3F686E174310 CRC64;

Query Match 63.3%; Score 38; DB 2; Length 986;
 Best Local Similarity 87.5%; Pred. No. 2.1e+02;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 RARVVWAA 10
 | | | | |
 Db 810 RARVVWAS 817

RESULT 14
 Q9YAA3 PRELIMINARY; PRT; 108 AA.
 ID Q9YAA3
 AC Q9YAA3;
 DT 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE HYPOTHETICAL 11.4 KDA PROTEIN APE2036.
 GN APE2036.
 OS Aeropyrum pernix.
 CC Archaea; Crenarchaeota; Desulfurococcales; Desulfurococcaceae;
 OC Aeropyrum.
 OX NCBI_TaxID=56636;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K1;
 RX MEDLINE=93310339; PubMed=10382966;
 RA Kwarabavasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
 RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
 RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
 RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
 RA Yamazaki J., Kishida N., Oguchi A., Aoki K.-I., Kubota K.,
 RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
 RT "Complete genome sequence of an aerobic hyper-thermophilic
 crenarchaeon, Aeropyrum pernix K1";
 RL DNA Res. 6:83-101(1999).
 DR EMBL; AP000063; BAA81046.1; -;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 108 AA; 11428 MW; FB81D7DF5FC4BAC8 CRC64;

Query Match 61.7%; Score 37; DB 1; Length 108;
 Best Local Similarity 66.7%; Pred. No. 33;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ARRARRVWAA 9
 | | | | |
 Db 63 AERARRVWVS 71

RESULT 15

O9KZA5
ID O9KZA5 PRELIMINARY; PRT; 256 AA.
AC O9KZA5;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE PUTATIVE SHORT-CHAIN DEHYDROGENASE.
GN SC5G8.09C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-A3(2);
RC Saunders D.C., Harris D.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN-A3(2);
RC Bentley S.D., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA STRAIN-A3(2);
RC MEDLINE-97000351; PubMed-8843436;
RA Redenbach M., Kieser H.M., Denapaita D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
CC -!- SIMILARITY: TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES (SDR)
FAMILY
DR EMBL; AL353872; CAB89060.1; -;
DR InterPro; IPR002198; adh_short.
DR Pfam; PF00106; adh_short; 1.
DR PROSITE; PS00061; ADH_SHORT; UNKNOWN_1.
KW Oxidoreductase.
SQ SEQUENCE 256 AA; 26809 MW; 8EFF5139F44D8BFA CRC64;

Query Match 61.7%; Score 37; DB 2; Length 256;
Best Local Similarity 70.0%; Pred. NO. 79;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ARARVVWAA 10
DB 35 ARRTVTWLA 44

Search completed: December 31, 2001, 08:53:35
Job time: 89 sec